GTN-Québec
Groupe de travail québécois sur les normes et standards en TI pour l'apprentissage, l'éducation et la formation

Publication 2009-03

Concept and prototype of an Aggregator Portal for Learning Opportunities Based on the MLO-AD Standard

---

Katharina Bauer-Öppinger

Mission du GTN-Québec

La mission du Groupe de travail québécois sur les normes et standards pour l'apprentissage, l'éducation et la formation (GTN-Québec) est de fournir une expertise à la communauté éducative en matière de normalisation.

Les membres du GTN-Québec proviennent des trois ordres d'enseignement, des ministères, ainsi que du secteur privé de la formation. En s'appuyant sur les travaux des groupes internationaux d'élaboration des normes, ils soutiennent les acteurs du milieu de l'éducation pour favoriser l'implantation de pratiques communes de description et de production de ressources éducatives interopérables, réutilisables et accessibles à tous.

Ces ressources forment un patrimoine éducatif d'une valeur inestimable pour les communautés éducatives francophones. Assurer son enrichissement et sa pérennité est en conséquence, depuis sa fondation, au cœur des préoccupations du GTN-Québec.

Objectifs du GTN-Québec

1. Dans une perspective d'accompagnement, consulter les acteurs du milieu de l'éducation pour mieux définir comment les approches basées sur les normes et standards peuvent aider à concrétiser la mission éducative de leur organisation ;

2. Connaître des solutions basées sur des normes et standards, s'assurer qu'elles correspondent à la réalité et aux besoins du milieu et proposer, le cas échéant, des adaptations ou des guides d'utilisation de ces normes;

3. Faire connaître et encourager les pratiques normalisées de production et de description de ressources éducatives ;

4. Favoriser le développement d'une masse critique de REA numériques accessibles, pérennes et réutilisables au sein des établissements de chaque ordre d'enseignement ;

5. Maintenir l'expertise et la représentation québécoises en matière de développement de normes internationales et d'autres standards.

Les activités du GTN-Québec sont réalisées avec l'appui financier du ministère de l'Éducation, du Loisir et du Sport du Québec et grâce à la collaboration de ses membres.

www.gtn-quebec.org

ISBN 978-2-924168-04-2 (PDF)

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2012
Dépôt légal – Bibliothèque et Archives Canada, 2012

Concept and Prototype of an Aggregator Portal for Learning Opportunities Based on the MLO-AD Standard

Katharina Bauer-Öppinger

DIPLOMARBEIT

eingereicht am
Fachhochschul-Diplomstudiengang

Engineering für Computer-Basiertes Lernen

in Hagenberg

im August 2009

Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und die aus anderen Quellen entnommenen Stellen als solche gekennzeichnet habe.

Hagenberg, am 30. Juli 2009

Katharina Bauer-Öppinger

Contents

Foreword

From August to December 2008, I worked at the Computer Research Center of Montréal (CRIM) for the mandatory internship of my studies. CRIM is a supporting member of GTN-Québec, which works in the field of e-Learning object standards for Québec and the Francophone world. Both CRIM and GTN-Québec are interested in analyses and further development of the Metadata for Learning Opportunity Advertising (MLO-AD) standard that specifies a data model for describing educational events. Therefore, Sacha Leprêtre (from CRIM) offered me a project for my diploma thesis about MLO-AD. Together with Yves Otis (from GTN-Québec) and Sacha, the idea of an aggregator portal for learning opportunities based on MLO-AD was developed to demonstrate the potential and practicability of this new standard.

At this point, I want to thank CRIM for giving me the opportunity to do my internship there and for allowing me to do this project for my thesis. Special thanks go to Sacha Leprêtre for his great support and impulse to develop additional ideas. I am also grateful to Yves Otis for being available for discussions about MLO-AD and related technologies.

Further thanks go to A.Univ.Prof. DI Dr. Wolfgang Schreiner for being my thesis adviser and providing me with valuable tips and suggestions concerning the improvement of this thesis.

Moreover, I am very thankful to my family, who made it possible for me to study at the University of Applied Sciences in Hagenberg and supported me, together with my friends throughout my whole studies and thesis work.

Abstract

The Internet provides much information about learning opportunities, including courses at universities, conferences, and any other types of training programs. Often, this information is cluttered, which increases the need for organizing and accessing the information in a way that is useful, structured, and comprehensive. The Metadata for Learning Opportunities Advertising (MLO-AD) standard specifies the data model for learning opportunities. It defines which information an educational event must or can provide to make it comparable with other programs. This thesis describes a concept for an aggregator portal which collects data based on the MLO-AD standard from multiple databases and consolidates it. Moreover, the portal allows information about learning opportunities to be accessed in a comprehensive and structured way, and keeps the data up-to-date. The emphasis of this concept is placed on different solutions of data collection that use the complex structure of MLO-AD. Besides the analysis for collecting the required information and the exploration of the design for such a portal, this thesis results in an implementation of a prototype. This prototype is the first MLO-AD-based web application which allows prospective learners to search for an educational event and providers of learning opportunities to advertise their offers.

Kurzfassung

Das Internet bietet zahlreiche Informationen über Lernmöglichkeiten. Dieser Begriff schließt Universitätskurse, Lernseminare und -konferenzen, sowie alle weiteren Arten von Lehrveranstaltungen ein. Häufig jedoch sind diese Informationen verstreut, was den Bedarf an einen nützlichen, strukturierten und umfassenden Zugang zu diesen erhöht. Der Metadata for Learning Opportunities Advertising (MLO-AD) Standard definiert das Datenmodell für Lernmöglichkeiten. MLO-AD bestimmt somit, welche Informationen von einer Lehrveranstaltung bereit gestellt werden können oder müssen, damit diese miteinander vergleichbar sind. Diese Diplomarbeit beschreibt das Konzept für ein Aggregator Portal, das Daten basierend auf dem MLO-AD Standard von mehreren Datenbanken sammelt. Zudem ermöglicht das Portal einen Zugang zu den strukturierten und umfassenden Informationen von Lernmöglichkeiten und aktualisiert diese regelmäßig, um sie auf dem neuesten Stand zu halten. Der Schwerpunkt dieser Diplomarbeit liegt auf der Analyse von unterschiedlichen Datenerfassungsmöglichkeiten, wobei die Daten die komplexe Struktur von MLO-AD verwenden. Neben dieser Untersuchung und der Designerforschung eines solchen Systems umfasst diese Arbeit auch die Implementierung eines Prototyps. Dieser Prototyp ist die erste Webanwendung basierend auf MLO-AD, die zukünftigen Lernenden die Suche nach Lehrveranstaltungen und Anbietern von Lernmöglichkeiten die Werbung ihrer Angebote ermöglicht.

Chapter 1

Introduction

Learning is not a product of schooling,
but the lifelong attempt to acquire it.

Albert Einstein (1879–1955)

Lifelong learning is important for many people due to the wish of knowledge enrichment for private or professional reasons, but also the need for employability and competitiveness. With the increasing demand for educational offers and the progressive development of technologies that can be used for education, the supply of learning opportunities has been rising.

> In this thesis, the term *learning opportunity* is used as a reference to all events that provide a context of learning. This includes inter alia courses at schools and universities, forums, conferences, as well as all kinds of training programs.

Because of the increasing number of learning opportunities, the rising availability of borderless education (anytime, anywhere), as well as the trend towards more mobility in education [23, p. 142], learners have more possibilities to choose their courses. However, this wide range of opportunities also makes it more difficult to get a survey of all offers.

To get information about learning opportunities, many learners use educational portals directly from a specific educational organization or from a provider which collects data from more institutions. Systems that gather data from multiple sites and consolidate it onto one page are called *aggregators*.

> Aggregation is the process of creating a compound object from several smaller ones [19, p. 4].

Usually, the decision for attending a course or any other educational program is preceded by retrieving at least the main information about this educational event as, for example, subject, time, location, costs, qualification, and level or quality of course. To ensure the provision of these properties and, more importantly, to be able to compare learning opportunities, a common structure for the data of educational offers must be defined. Uniform models of data or any technology are guaranteed by standards or specifications.

In December 2007, the Metadata for Learning Opportunity (MLO) project group started to develop the MLO-Advertising (MLO-AD) standard that addresses metadata sufficient for advertising a learning opportunity [39, p. 2]. MLO-AD provides learners with information about an educational event and helps them to compare different offers. Since October 2008, MLO-AD has been in development as a formal European standard. Also in Canada, a group is contributing to the development of MLO-AD. The working group GTN-Québec[1] supplies the educational community with expertise in the area of e-Learning object standards and aims to propose an application profile based on MLO-AD. Different interest groups contribute to the targets of GTN-Québec. One of them is the Computer Research Center of Montréal (CRIM)[2]. CRIM is an information technologies-applied research center that develops and transfers technologies and knowledge. Additionally, it provides a training center for information technology. As CRIM and GTN-Québec are generally interested in the developments of MLO-AD, these organizations made the project of this thesis possible and have strongly supported the creation of this work.

This thesis describes the concept of an aggregator portal that provides information about learning opportunities, which are structured by the data model of MLO-AD. Thereby, the emphasis of the analysis is set on possibilities for collecting information from different learning opportunity providers, which include all institutions and instructors offering educational events. The challenge of this research is posed by the complex structure of MLO-AD that makes the collection and transfer of such data more difficult. Besides the concept of the portal, this work includes the implementation of a prototype. This prototype is the first aggregator portal based on the MLO-AD standard that provides a search functionality and the possibility to advertise learning opportunities.

The structure of this work is as follows: Chapter 2 gives a review of prior standards for describing learning opportunities, the MLO-AD standard itself, existing portals that offer search functionality for educational events, and existing aggregator systems. Chapter 3 describes the target audience and the requirements of the aggregator portal based on MLO-AD, as well as the

[1] http://www.gtn-quebec.org
[2] http://www.crim.ca

criteria for the prototype. Different technologies for collecting data about learning opportunities are analyzed in Chapter 4, while the emphasis is set on solutions for the complex structure of MLO-AD. Chapter 5 addresses the user interface and the possibilities for interaction with the aggregator portal. The implementation of the web application, which collects data from learning opportunity providers, is described in Chapter 6 and reflected in Chapter 7. The latter also evaluates the analyses of this thesis. Appendix A describes and lists the data saved on the enclosed CD-ROM.

Chapter 2

State of the Art

MLO-AD is not the first standard that offers a common structure to describe and compare learning opportunities. Different specifications across Europe and in the USA have already been used for several years. Some of them also build the basis of web applications that allow users to search for information about learning opportunities.

This chapter evaluates existing standards for describing learning opportunities to gain a better insight into the development of MLO-AD, as well as the connections and influences of those standards among each other. The history, the structure, and the specialties of MLO-AD are likewise addressed. One following section also takes a closer look at some available web portals that offer search functionality for educational events. To give an insight into the system of aggregators, representative data collectors are additionally described in this chapter.

2.1 Existing Standards

Universities or other educational institutions have used various data formats, data models or terminology to describe their offers in an electronic way. This makes a comparison of learning opportunities and interoperability between portals of these institutions very difficult or impossible. Non-interoperability of data is usually caused by the lack of widely accepted standards or specifications.

Across Europe, a trend to develop comparable descriptions of educational offers arose after the Bologna Process, which was initiated in June 1999. Today, 46 countries are committed to achieving the goals of this process and

to creating the European Higher Education Area (EHEA)[1]. The Bologna Declaration of June 19, 1999 says [13, p. 7]:

> The achievement of greater compatibility and comparability of the systems of higher education nevertheless requires continual momentum in order to be fully accomplished. We need to support it through promoting concrete measures to achieve tangible forward steps.

Therefore, European governments have strongly encouraged projects which targeted standardization of course descriptions to achieve a better comparison of educational offers and interoperability between different educational systems.

2.1.1 CDM and CDM-fr in Norway and France

The specification *Course Description Metadata* (CDM)[2] was developed in 2001 by USIT's XML group[3] at the University of Oslo in order to make study programs or course units comparable and university portals interoperable. CDM is a part of the Norwegian e-Standard project, Norway Opening Universities (NOU)[4], a national initiative for change and innovation in Norwegian higher education. It is expressed as a W3C XML schema document[5] and has been adopted by all Norwegian universities as a base for Norway's national educational portal. CDM was presented in 2005 to the CEN/ISSS as a "candidate standard" for course description [30, p. 2]. The Information Society Standardization System (ISSS) is the Information and Communication Technologies sector of the European Committee for Standardization (CEN), which is a business facilitator in Europe, and provides a platform for the development of European standards and other technical specifications[6].

According to [20, p. 2], the concept of CDM is defined as follows:

> CDM addresses the description of educational course units or other forms of educational offerings at all levels. It specifies the structure and semantics of the key concepts used in course descriptions. The metadata are specified as an XML schema, and guidelines with examples are given to facilitate the generation

[1] http://www.ond.vlaanderen.be/hogeronderwijs/bologna
[2] http://cdm.utdanning.no/CDM
[3] http://www.usit.uio.no/saus/xml (in Norwegian)
[4] http://norgesuniversitetet.no/seksjoner/english
[5] http://www.w3.org/XML/Schema
[6] http://www.cen.eu/cenorm/businessdomains/businessdomains/isss

of course descriptions as XML documents. The metadata are intended to satisfy the following objectives:

- Facilitate description and exchange of information about educational course units,
- Facilitate standardization of course unit descriptions,
- Facilitate the establishment of national and international course catalogs,
- Facilitate the establishment of course portals and other services helping students.

Besides the information about the course content or study program, students need to know details related to the courses to be able to choose to study at a certain institution. Therefore, CDM not only lists and describes course units and their content, but also supplies further information needed by students [30, pp. 3–4]. The data format is divided into four main parts, as is apparent from the XML schema of CDM[7]:

- *Organization unit (orgUnitType)*: Information about the institution, like description, kind of institution, student facilities, regulations, admission.
- *Study program (programType)*: Description of a program comprising a set of course units. This concept includes, for example, name, description, qualification, level, prerequisites, teaching form, expenses, program structure and duration.
- *Course unit (courseType)*: Information on course content, degree, credits, level, syllabus, admission and prerequisites, teaching place, language and more.
- *Contact person (personType)*: Information on all the relevant data of the contact person.

CDM fully supports the European Credit Transfer and Accumulation System (ECTS) and Diploma Supplement (DS) for comparing the study attainment of students of higher education across Europe [20]. This means it covers all specified ECTS items which make study programs easy to understand and comparable for all students and staff — local and foreign. CDM is also compatible with a number of existing standard specifications, such as IEEE Learning Object Metadata (LOM) for course content description [30, p. 2].

[7] http://utdanning.no/schemas/CDM/2.0.4/CDM.xsd

CDM-fr

Within a project coordinated by the French Ministry of Education, French universities and other higher education institutions modified the CDM specification and developed CDM-fr[8]. The project group started in March 2004 with the goal of adjusting the XML schema of CDM to the needs of universities in France. CDM-fr keeps the four main parts of Norway's CDM for its XML schema to describe courses and study programs. However, CDM-fr has some new or modified specifications in the main parts of the schema, uses new vocabulary for France, and is extended with the domain "habilitation" for authorizations. A mind map on the website of CDM-fr[9] illustrates the structure of the modified XML schema. CDM-fr is accepted as an AFNOR[10] standard, which is the standards association of France.

Many universities all over France use CDM-fr for describing and advertising their courses and training programs. Information about these projects and the institutions are collected on the website of CDM-fr[11].

2.1.2 EMIL in Sweden

The history of the Swedish specification *Education Information Markup Language* (EMIL)[12] started in the same year as CDM — 2001. The National Agency of Higher Education developed this metadata model as an XML schema for their online service www.studera.nu to describe parts of the Swedish educational sector. In the beginning of 2003, a steering committee was constituted to build a national catalog containing information about all courses and programs within the publicly financed educational sector of Sweden. The committee was comprised of representatives from the National Agency for Education, the National Labor Market Board (AMS) and the National Agency for School Improvement (MSU) [27, p. 4]. This catalog is built on an information service that collects and distributes the course information files, as well as a common information model to describe the course information. EMIL is the chosen model for the national catalog. The goal of EMIL is to provide a metadata model for many different information producers like universities, colleges, folk high schools, and municipalities.

As described in [27, p. 8], the XML schema of EMIL is based on following three concepts:

[8] http://cdm-fr.fr (in French)
[9] http://cdm-fr.fr/ressources/cartes-heuristiques-cdm-fr/cartes-cdmf
[10] http://www.afnor.org/portail.asp?lang=English
[11] http://cdm-fr.fr/ressources/ressources/carte-cdmfr-France (in French)
[12] http://mjukis.skolutveckling.se (in Swedish)

- *EducationProvider*: Describes the provider of the education based on the vCard standard (RFC 2426). This concept contains the school's name and address, as well as information about the contact person for each particular provider.
- *EducationInfo*: Contains general information about a certain course or education. The concept includes, e.g., course name and course description.
- *EducationEvent*: Describes a certain education event, such as the start date and application code. This concept refers to one EducationInfo and one EducationProvider.

EMIL's specification includes the definition of an "information hub," which is an information service that collects and distributes EMIL files and provides course information to various end-user services [27, pp. 4–5]. One advantage of an information hub is the allocation of a cache in the system between the providers and the end-user services for better accessibility and performance. Other reasons are to provide a common and neutral arena which all information providers can contribute to or to be able to cater to the needs of many services. An example for information services using EMIL is the SUSA hub run by the National Agency for Education and used by the end-user service Utbildningsinfo.se.

2.1.3 XCRI in the United Kingdom

The project *eXchanging Course-Related Information* (XCRI) began in April 2005 [34, p. 1]. It is funded by the Joint Information Systems Committee (JISC)[13] and operates in partnership between Manchester Metropolitan University, the JISC-CETIS service at the University of Bolton, and KaiNao Ltd. It is specially developed for the educational market of the United Kingdom.

The core development of XCRI is the XML specification XCRI Course Advertising Profile (XCRI-CAP), which describes course-related information that encompasses course quality assurance, course marketing, course enrolment and requirements[14]. The XCRI-CAP format allows learning providers to publish their course information, so that it can be easily collected by organizations with course search services such as UCAS[15] (UK's national course information and applications aggregator).

Before developing the XCRI Course Advertising Profile, the project group reviewed existing course information standards to assess their suitability for

[13] http://www.jisc.ac.uk
[14] http://www.xcri.org
[15] http://www.ucas.ac.uk

the needs of the UK [34, pp. 4 and 6]. The most promising specification was the XML schema of the Norwegian CDM. The decision to build on the CDM work was also based on its particular attention to the European Credit Transfer Scheme (ECTS). However, some structural problems which compromised CDM's ability to fulfill the needs of UK's educational systems were identified:

- In CDM, the course specification and its offering are conflated. However, several UK institutions need a separation of these two elements, as a specification can be offered many times or in multiple locations.
- CDM's hierarchy of programs and courses is too strict for the UK's educational institutions. There would be a need for more flexibility, as, for example, a possibility to add more elements between CDM's program and course objects.

Besides CDM, other standards, like the Swedish EMIL specification, were also analyzed. EMIL incorporates the missing features of CDM, but it is not well applicable for realizing XCRI's objectives of addressing quality, marketing, reporting and enrolment requirements [34, p. 4]. Therefore, XCRI decided to develop its own specification. Nevertheless, XCRI still cooperates with the CDM and EMIL teams to advance understanding of curriculum metadata requirements and possibilities [30, p. 3].

XCRI-CAP provides following core elements, as shown in its XML schema[16]:

- *Catalog*: This element contains providers of courses or educational training programs.
- *Provider*: This element covers all relevant information about the institution that provides courses or other educational offers. A provider can also include one or several courses.
- *Course*: This element handles information about the course, such as title, description, qualification, and is comprised of presentations.
- *Presentation*: This element includes description, start, end, duration, study mode, costs, language, entry requirements, available places, and venues.
- *Venue*: This element covers description, address, phone, email, URL of the venue of a presentation.

[16] http://www.xcri.org/Tools.html

In addition to the national system UCAS, many universities started projects using XCRI to publish and compare their educational offers. Examples are collected on the website of JISC[17].

2.1.4 PAS 1068 in Germany

The *Publicly Available Specification* (PAS) 1068, which was published in January 2007, is a guideline for the description of educational offers. It is not subject to national, European or international standardization.

PAS 1068 is described in [14, pp. 3–4] in the following way:

> With this PAS, the "Transparency in e-Learning" working group at the German National Standards Body DIN[18] provides a description scheme that allows the providers to describe their educational offers in terms of a "leaflet." The minimum set of data is specified to standardize the description of educational offers and to enable their comparison. [...] The PAS is applicable to all processes in learning, education, and training, and particularly includes the consideration of e-Learning.

The description scheme defines criteria which are mandatory, optional or optionally mandatory (they have to be provided if they exist) and how they should be described (yes/no-answer or detailed information). PAS 1068 also uses references to the IEEE specification Learning Object Metadata (LOM) and to the German PAS 1045 (Further education and professional training databases and information systems — criteria of the contents and for data exchange formats). As PAS 1068 especially considers e-Learning, it offers many criteria concerning technical aspects, data recording, data processing, accessibility and functional aspects that are less available in the aforementioned specifications. With the information structured by the guideline of PAS 1068, providers cannot only describe their educational offers, but also promote them and make them comparable. The description scheme of PAS 1068 is usually constructed in form of a table, but it is also provided as an XML binding[19].

2.1.5 PESC in the USA

Not only within Europe does the harmonization and easier exchange of course information need to be achieved. In 1997, the *Postsecondary Electronic Stan-*

[17] http://www.jisc.org.uk/whatwedo/programmes/elearningcapital/courseinfo.aspx
[18] http://www.din.de
[19] http://www.qed-info.de/PAS (in German)

dards Council (PESC)[20] was founded in Washington, D.C. to lead the establishment and adoption of data exchange standards in education. Its goals are to enable the improvement of institutional performance and foster collaboration across educational communities in order to lower costs, improve service, and attain system interoperability.

PESC focuses on the development of electronic standards for student-related information, but it also considers information about courses in its specifications. The first standards in higher education were already developed before PESC in the early 1990's, and they were based on the EDI (electronic data interchange) transcript standards. These standards include the Educational Course Inventory (Transaction Set 188), which is used by postsecondary educational institutions to transmit course information. Nowadays, admissions, financial aid, and registrar communities are developing XML standards under PESC. In 2005, the development of a Course Catalog in XML was initialized to include institutional and curriculum information in the XML Postsecondary Transcript of PESC [37]. PESC also observes the development of the XCRI in the United Kingdom, and incorporates parts of it in its Course Catalog. However, the XML Course Inventory workgroup has been inactive since January 2009.

2.2 History and Development of MLO-AD

As described in Section 2.1, the existing specifications for providing information about learning opportunities are all national standards, and are not used across different countries.

In 2004, CEN/ISSS WS-LT[21] (Workshop on Learning Technologies at the European Committee for Standardization) proposed a project for the harmonization of the existing European specifications, which are CDM, CDM-fr, EMIL, XCRI and PAS 1068. Therefore, a European-wide standard should be created to describe learning opportunities.

After CDM did not get funding by the Directorate General (DG) Industry of CEN [38, p. 4], a group of experts in this field continued the development of this harmonization on a voluntary, unpaid basis. In November 2007, representatives from European and US universities and other institutions met in Rome to discuss future common data standards for exchanging student curriculum data. Requirements for harmonization of course description were summarized by a technical committee of CEN in the following so-called "Athens Declaration" [31]:

[20] http://www.pesc.org
[21] http://www.cen.eu/isss/Workshop/lt

- There is a considerable interest in many countries in Europe in creating specifications for the exchange of information about courses and other learning and training opportunities.
- There is a clear scope for greater harmonization of these efforts within a European context.
- All existing national initiatives will benefit from contributing towards harmonization at a European level.
- There are sufficient clear commonalities across existing national initiatives for future European standards to be developed.
- Harmonization should balance the benefits of common standardization with the necessity of meeting local contextual needs and infrastructure.
- Harmonization efforts should focus on small, simple models based upon existing commonalities that can be expanded upon at national or regional level, rather than all-inclusive, monolithic standards.

The Athens Declaration is also applicable to student curriculum data which cover all personal information, courses attended, and grades attained by the student.

In December 2007, the group of experts decided to build the new standard on the core elements of CDM and changed the name of the specification to *Metadata for Learning Opportunities* (MLO) [38, p. 7]. The focus of the project is the advertising of learning opportunities, and therefore, the development of the *MLO-Advertising* (MLO-AD) standard, which is described in [39, p. 2] as follows:

> MLO-AD is a standard addressing metadata sufficient for advertising a learning opportunity. The goal of MLO-AD is to provide information about a learning opportunity, to enable the learner to make a decision if there is a need for more information about the learning opportunity, and where to find that information. The group also aims at developing a lightweight standard which is designed to facilitate semantic technologies and web architectures to support several mechanisms for exchange of the information and aggregation of information by third party service suppliers.

CEN endorsed MLO-AD as a CEN Workshop Agreement (CWA) in October 2008[22] [23] and committed itself to the further development of the specification

[22] http://www.cen.eu/cenorm/standards_drafts/index.asp
[23] http://zope.cetis.ac.uk/members/scott/blogview?entry=20081021140752

into a formal European standard known as the European Norm (EN). After the agreement upon an EN (this development may take up to two years), MLO-AD will be a standard in all 30 member countries of the CEN.

MLO-AD should not be limited to Europe. GTN-Québec[24] is a working group whose mission is to provide the educational community with expertise in the area of e-Learning object standards in order to promote the creation and enrichment of an educational heritage for Québec's education community, as well as that of the Francophone world. GTN-Québec aims to propose an application profile based on the MLO-AD specification to the community. Additionally, it provides working examples of XML binding and RDFa/microformat, as well as implementation guidelines. Different interest groups collaborate to reach the targets of GTN-Québec. One of them is the Computer Research Center of Montréal (CRIM)[25], which also provides services for education and training. The CRIM Training Center aims to fulfill the training needs of businesses and organizations in the field of information technology[26].

MLO-AD presents an abstract model for representing and advertising learning opportunities. The data model is based on the object of the learning opportunity and specifies three resources, which are the provider, the specification and the instance of the learning opportunity. The objects of MLO-AD are described in [39, pp. 6–7] in the following way:

- *Learning Opportunity (LO)*: A chance to participate in education or training.
- *Learning Opportunity Provider (LOP)*: An agent (person or organization) that provides learning opportunities.
- *Learning Opportunity Specification (LOS)*: An abstract description of a learning opportunity, consisting of information that will be consistent across multiple instances of the learning opportunity.
- *Learning Opportunity Instance (LOI)*: A single occurrence of a learning opportunity. Unlike a Learning Opportunity Specification, a Learning Opportunity Instance is not abstract, may be bound to particular dates or locations, and may be applied for or participated in by learners.

Thereby, the following issue must be regarded as described in [39, p. 3]:

> The model proposed within the standard is not intended to define the electronic representation of learning objects in general — the

[24] http://www.gtn-quebec.org
[25] http://www.crim.ca
[26] http://www.crim.ca/en/services/Formation

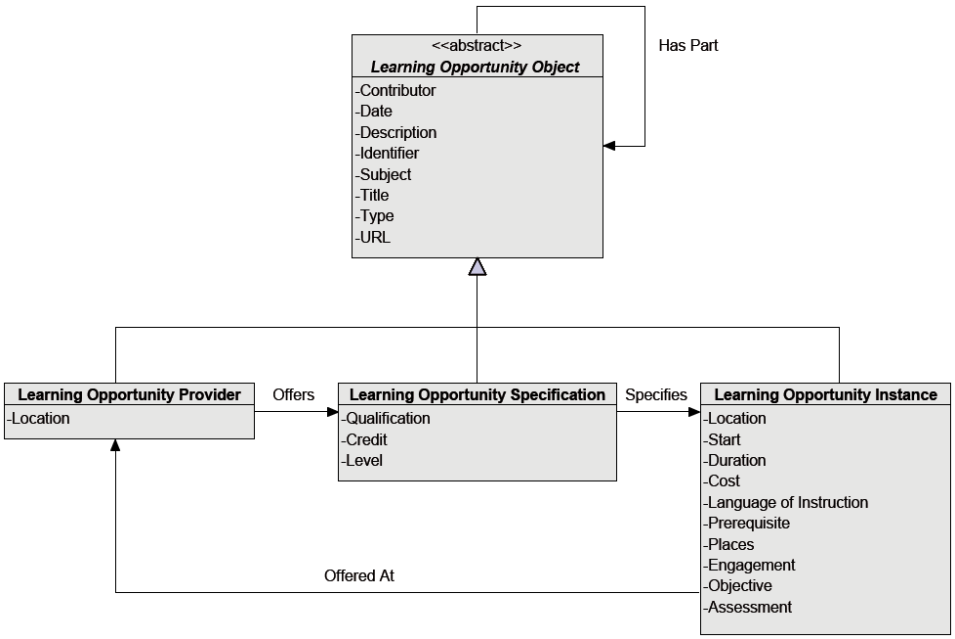


Figure 2.1: Illustration of the MLO domain model (according to [39, p. 6])

> scope of the standard is restricted to define the electronic representations of learning opportunities to facilitate their advertising and subsequent discovery by learners.

The model of the MLO and the associations of its objects are illustrated on Fig. 2.1. During the creation of the model design, attention was also drawn to supporting the ECTS description and the exchange of ECTS information. The standard only describes the model, and does not address the vocabularies needed to ensure semantic interoperability between different educational and jurisdictional domains. The reason for not dealing with vocabularies is based on the need for frequently updating and maintaining the vocabularies. Therefore, all vocabularies will be maintained as separate CEN Workshop Agreements (CWAs) by the CEN/ISSS WS-LT [39, pp. 2–3].

Further developments of the MLO are specified in [39, p. 3] as follows:

> In the future, the MLO set of standards will be further developed to describe Metadata for Learning Opportunities related to the Europass system used throughout Europe. Based on other needs for metadata related to Learning Opportunities, new standardization projects [besides the project for advertising learning opportunities; author's note] could also be launched.

2.3 Existing Portals for Learning Opportunities

Often, learners want to get information about learning opportunities from several educational institutions. Therefore, many use portals which collect course and program data from multiple institutions and present them in a comprehensive and structured way. Some platforms which allow learners to search for learning opportunities across one or more countries already exist. Examples of sites which help to find a course or training program are SchoolFinder.com[27], UCAS in the UK[28], FindaCourse.com[29], Ploteus[30] (an EU project), and FastTomato.com[31]. This section describes three representative portals for learning opportunities more precisely.

2.3.1 Moveonnet

Unisolution[32] is a software and consulting company for institutions of higher education, founded by students of the Technical University Darmstadt in Germany. In 2006, it launched the web portal *moveonnet*[33], which provides a comprehensive directory of higher education worldwide. According to the information on its homepage, moveonnet is especially relevant for international relations officers, international students, and international researchers.

The base of the portal of moveonnet is the *Worldwide Directory of Higher Education*, which provides the information about higher education in four areas[34]:

- Institutions of higher education, including general information, contacts, partner lists, information for exchange students, ranking positions, map locations, and more.
- Countries, including general information, regions/states, higher education system, institution types and list of institutions.
- Networks, including general information, aims, contacts and members.
- Intensive courses, such as summer courses or language courses, including description, modalities, and contacts.

[27] http://schoolfinder.com
[28] http://www.ucas.ac.uk
[29] http://findacourse.com
[30] http://ec.europa.eu/ploteus
[31] http://www.fasttomato.com
[32] http://www.unisolution.eu
[33] http://www.moveonnet.eu
[34] http://www.moveonnet.eu/about

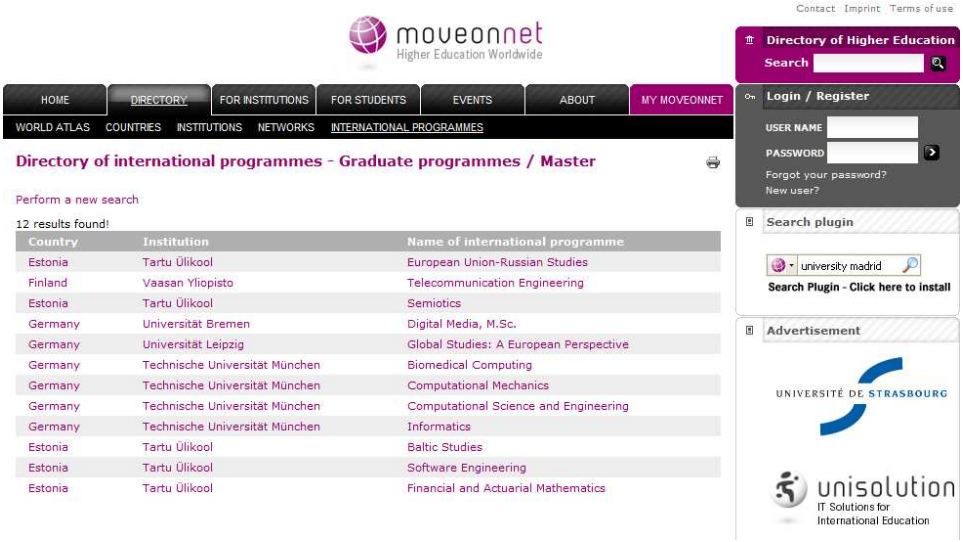


Figure 2.2: Screenshot with search results of moveonnet's international programs

The user interface for searching in the worldwide directory for international programs includes input and select fields for the keyword, the program type (Bachelor, Master, PhD, or summer school), and the desired country. Fig. 2.2 shows the results of a search for Master programs in Computer Science.

Each institution is presented to students and partners in a standardized and comparable way. An institution can register onto moveonnet without a registration fee. The institution itself is responsible for the up-to-dateness of its data and can enter and update the information anytime, free of charge. More than 1,000 institutions of higher education are already registered on moveonnet.

Moveonnet also provides a Search Plugin, which can be installed for the web browser to directly search the worldwide directory of higher education without accessing the website.

2.3.2 CourseAtlas and CollegeTransfer

The company AcademyOne[35], located in Pennsylvania, United States, was founded in 2005 with the target to enable new forms of academic collaboration on the Internet. "We are dedicated to improving the underlying infrastructure, applications and processes supporting the diverse array of educational institutions while reducing redundancy, complexity and cost," says

[35] http://www.academyone.com

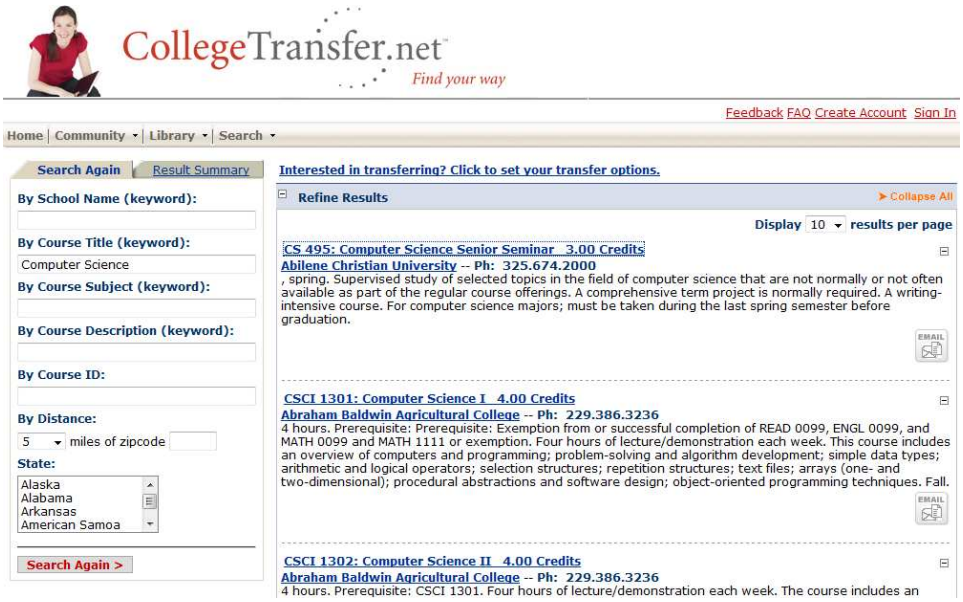


Figure 2.3: Screenshot of CollegeTransfer, showing search results with the course title "Computer Science"

David K. Moldoff, founder of AcademyOne, on the company's web page[36]. AcademyOne offers two web-based platforms for students: *CourseAtlas*[37] and *CollegeTransfer*[38].

CourseAtlas is a searchable course database with more than three million courses offered by over 4,000 colleges and universities within the United States. Courses can be searched free of charge by ID, title, description, subject, school name or location. Once a year, AcademyOne updates and aggregates the information about courses of colleges and universities. However, institutions can submit their course catalogs to AcademyOne anytime to have them included in CourseAtlas.

After two years of research, CollegeTransfer was launched in 2007 to help all involved parties in the process of student and academic credit transfer from one to another college or university. This comprehensive portal offers possibilities for students to evaluate transfer options and to save and share information about them, their attended courses, and their credits in a secure way. Institutions of higher education can use CollegeTransfer to manage the workflow of a college transfer, to build partnerships, and to advertise and share their course information. It also provides information about how aca-

[36] http://www.academyone.com/AboutUs/OurMission/tabid/199/Default.aspx
[37] http://www.courseatlas.com
[38] http://collegetransfer.net

demic credits can be evaluated to be acceptable for different educational providers. CollegeTransfer also uses the database of CourseAtlas to facilitate the search of courses and schools. Fig. 2.3 shows the user interface of CourseAtlas in CollegeTransfer and the result of a search by keywords.

AcademyOne uses PESC standards for its platforms to collect and display information about courses and students.

2.3.3 Hotcourses

The history of the *Hotcourses*[39] course guide began in 1996 in paper form, when the Hotcourses magazine was published in London. In 2000, the online version Hotcourses.com followed, offering information about 50,000 courses. Nowadays, the database covers information about more than one million courses from more than 17,000 providers within the United Kingdom. Therefore, Hotcourses can be designated as UK's largest course finder. Besides its headquarters in London, Hotcourses opened offices in India[40] (responsible for Uniguru.com — a study abroad site for Indian students, offering information of over 350,000 courses), Australia[41] (more than 45,000 courses) and South Africa[42] (over 2,000 courses).

Hotcourses offers a simple search mask which is permanently available on the website. The user can enter a subject keyword, course type (part-time, full-time without degree, undergraduate, postgraduate, MBA, home study), town or postcode. After the results of a search query have been displayed, the search can be further refined by defining a preferred qualification, duration or hours of study. Fig. 2.4 shows the results of a search for postgraduate medicine courses.

Providers of learning opportunities can sign up on Hotcourses.com for free. It is also in their hands to update the information about their courses and to ensure that their listings remain accurate. Every week, Hotcourses collects data of new courses to add them to the site. Information about undergraduate courses is provided by the UCAS (UK's national course aggregator), which uses the XCRI standard.

[39] http://www.hotcourses.com
[40] http://www.hotcourses.co.in
[41] http://www.hotcourses.com.au
[42] http://www.hotcourses.co.za

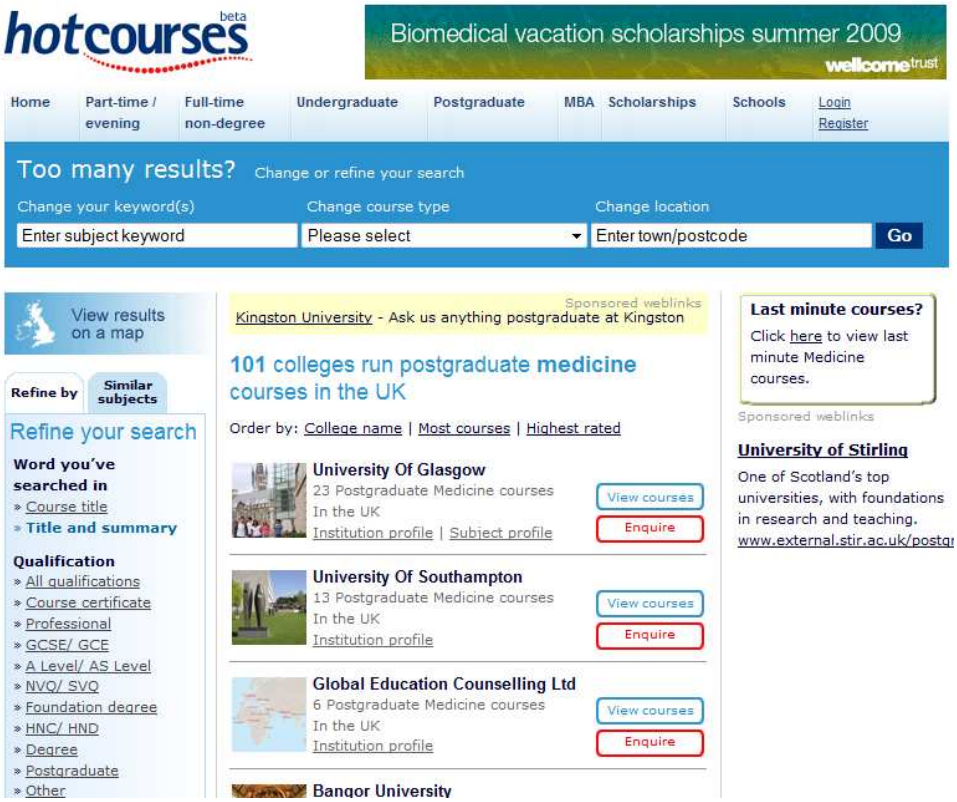


Figure 2.4: Screenshot of Hotcourses with the search result of postgraduate medicine courses in the UK

2.4 Existing Aggregators

A challenge in organizing a database is to keep the data up-to-date at all times and to gather recently emerged information. To solve this problem, it needs to have an automatic aggregator system which harvests data from its sources at regular intervals.

Aggregators have been used in many different applications. One example is an application which collects bank account data, investment account data, and other account data on one page. Nearly every bank that offers e-Banking supports this feature, and consolidates all the different accounts of a client.

Sites that provide price comparisons of different suppliers also use aggregation functionality to collect this data. Examples for these web applications are PriceGrabber.com, PriceRunner.co.uk, Shopbot.ca, and Geizhals.at.

Certain e-mail scanning software, like Microsoft Outlook and Mozilla Thunderbird, can act as "e-mail aggregators," because they show newly received

messages from multiple mailboxes of a user. Consequently, the user does not need to check each mailbox separately.

Another category of aggregators are "media aggregators," which can automatically download video and audio files due to "podcasts." Podcasts are special types of RSS feeds, where RSS is a standardized format to publish frequently updated content [42, p. 78] (Therefore, these aggregators are also called "Podcatchers"). Examples for media aggregators are Apple iTunes, Nullsoft Winamp, and Microsoft Zune.

Three of those aggregator systems that collect data regularly and automatically are described in this section more specifically.

2.4.1 Google News

In 2002[43], Google launched the News Aggregator *Google News*[44], which is described on its homepage in the following way:

> Google News is a computer-generated news site that aggregates headlines from more than 4,500 English-language news sources worldwide, groups similar stories together and displays them according to each reader's personalized interests.

Fig. 2.5 is a screenshot of its homepage, showing the main articles and headlines, as well as their sources. Google News supports many different languages and provides more than 40 regional editions. In general, Google News aims to promote original journalism. This is achieved by harvesting from professional news sites. The selection and ranking of the articles in Google News are totally automated and, therefore, not influenced by human editors. The ranking is mainly based on following factors[45]:

- Freshness of content
- Diversity of content
- Rich textual content which would help users searching for information to find the articles

The only human editorial input into the system is the list of sources which Google News harvests the articles from. However, Google accepts URLs of other news sites if a user wishes to include them in Google News. Users can

[43] http://googleblog.blogspot.com/2006/01/and-now-news.html
[44] http://news.google.com
[45] http://www.google.com/support/news_pub

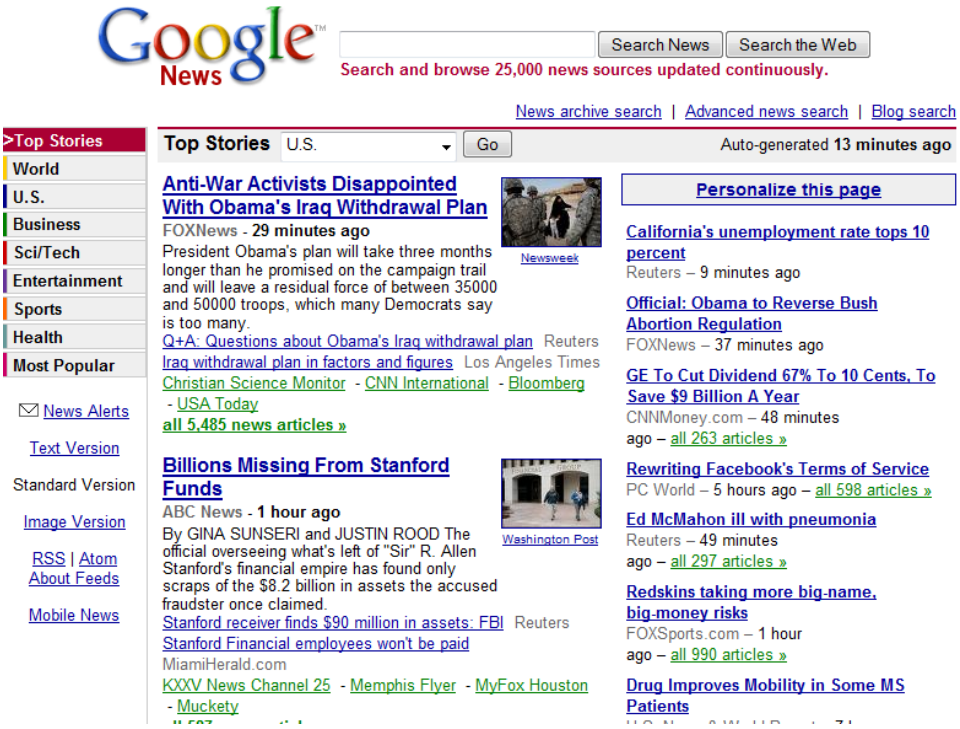


Figure 2.5: Screenshot of Google News

customize their Google News site and choose preferred sections, their locations and the number of displayed articles. Google also provides an easier access to updates about favorite topics via RSS and Atom feeds. After subscribing to a Google News feed, users will regularly receive a summary of new articles they are interested in.

2.4.2 We Feel Fine

The independent artwork *We Feel Fine*[46] was developed by Jonathan Harris and Sepandar Kamvar with the target to create "an exploration of human emotion on a global scale." Since 2005, their data collection engine has automatically harvested human feelings from a large number of weblogs every ten minutes. Therefore, the system searches the entries of weblogs for the occurrences of the phrases "I feel" and "I am feeling" and stores the found sentence in its database. After that, it identifies the "feeling" in that sentence (for example, happy, sad, angry) with a provided list of about 5,000 pre-identified "feelings." Moreover, the system saves the time the blog en-

[46] http://wefeelfine.org

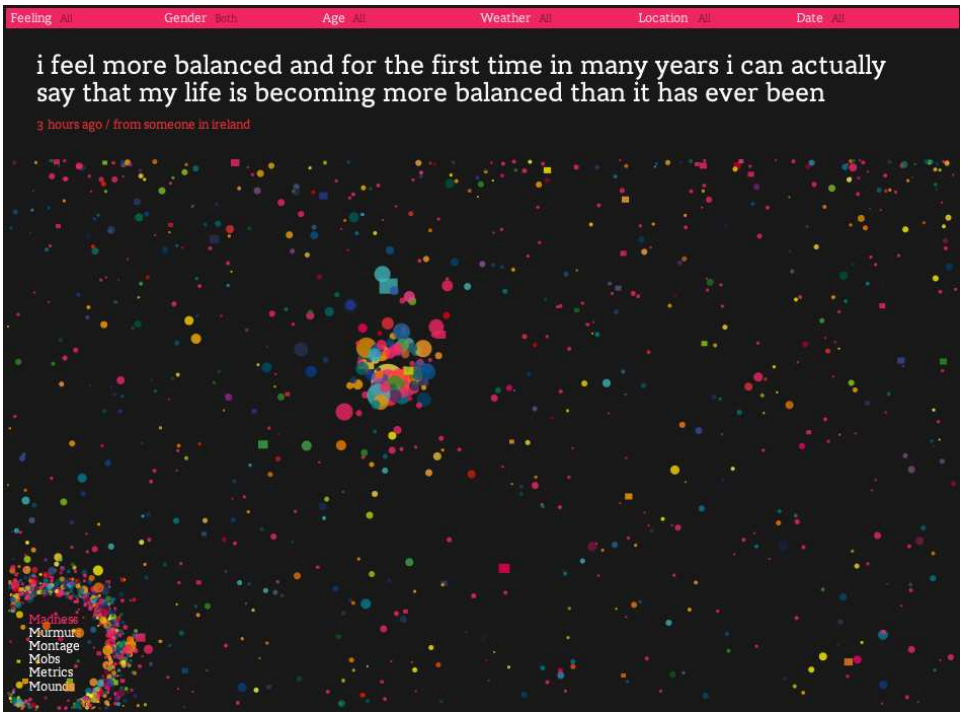


Figure 2.6: Screenshot of the We Feel Fine applet in the so-called "Madness" movement

try was written, the geographical location, the age, and the gender of the author, if this information is provided or can be extracted from the blog system. Even the local weather conditions are calculated from the data of time and location. If the blog entry also contains an image, this image will be saved along with the sentence. Accordingly, the aggregator system of We Feel Fine collects 15,000 to 20,000 new feelings per day. This data can be queried with the We Feel Fine applet in six different statistical movements. One of these movements is called "Madness," which was designed to show the feelings of the human world from a bird's eye view. This movement is illustrated in Fig. 2.6.

Jonathan Harris and Sep Kamvar created the data collection engine of We Feel Fine with Java, Perl, MySQL and Apache. The Processing software[47] by Ben Fry and Casey Reas is used for the applet. The code of We Feel Fine is closed source. However, the data of the collected feelings are freely available through the public API of We Feel Fine.

[47] http://www.processing.org

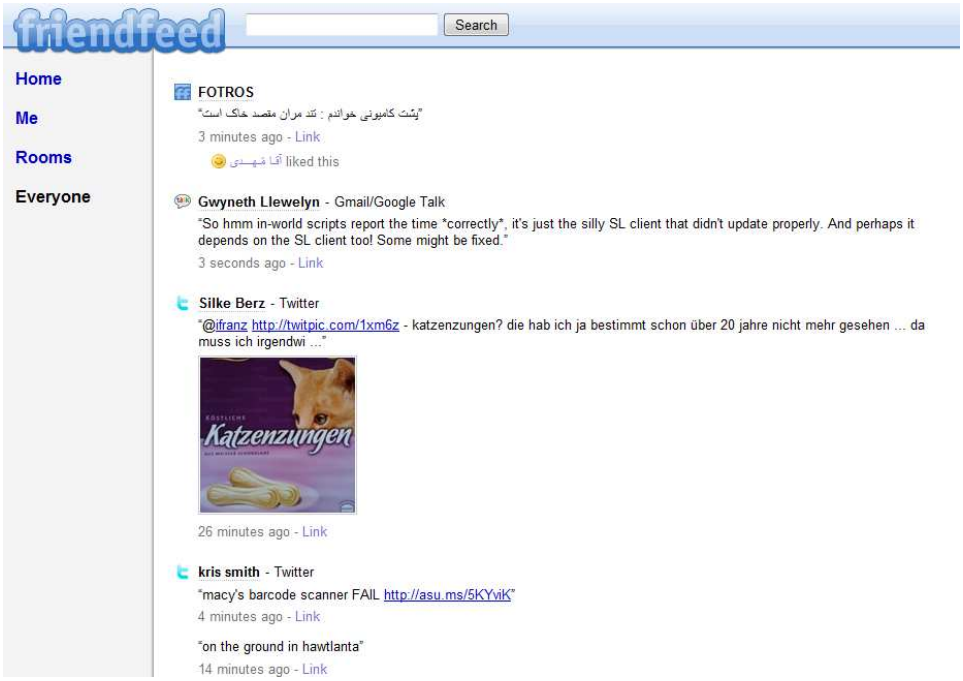


Figure 2.7: Screenshot of FriendFeed showing information shared by everyone

2.4.3 FriendFeed

The web application *FriendFeed*[48] collects shared information of users and their friends from different sites the users have chosen to aggregate. Therefore, the users are always up-to-date on web pages, photos, videos and music that their families and friends are sharing. FriendFeed also provides possibilities to leave comments on shared information or to start discussions among friends.

More than 40 sites are supported by FriendFeed, e.g., Amazon.com, Dailymotion, Facebook, Flickr, Gmail, Google Talk, Mister Wong, SlideShare, Twitter, YouTube, and blogs. The user usually just has to provide the username of a site; FriendFeed automatically finds and aggregates the public activity of this site using web crawling technologies similar to those of search engines. For each user, FriendFeed creates a web feed customized to the content the user's friends shared. This feed includes links to the sources of the shared information.

FriendFeed also offers rooms for sharing and discussing content concerning a special topic. Users have the possibility to create their own rooms for a

[48] https://friendfeed.com

specific topic or to allow just a few friends to see their shared information. Content can also be public in FriendFeed, which means that it is shared with everyone. Fig. 2.7 shows a screenshot of FriendFeed displaying publicly shared information.

With the provided API of FriendFeed, developers can interact with the website and use the shared information. They can also develop an interface for a mobile device or integrate FriendFeed into a web application.

Chapter 3

Requirements on the MLO-AD Aggregator Portal

The majority of the existing portals allowing a comparison of learning opportunities lists courses or programs, but does not provide much information about other learning events, such as forums or conferences. Except for the specification of time and location for this kind of training, further details are usually not available.

Moreover, existing portals often consider learning opportunities within one country only. This is due to the fact that a comparison of offers from multiple countries is more difficult by reason of different educational systems and various data formats that are used by training institutions for describing their offers.

The MLO-AD standard ensures a harmonization of data models describing learning opportunities across all countries that accepted this standard (see Section 2.2). MLO-AD defines which information must and can be provided about each opportunity to make them comparable and to provide comprehensive information for the user. Therefore, the need for an international comparison of all kinds of learning opportunities can be fulfilled by a system based on the MLO-AD standard.

This system must not only collect and provide the information about learning opportunities, but also keep it up-to-date. For example, subjects and contents of courses or training programs can be revised, starting dates can be updated for every new session or costs can be adjusted. Therefore, data of learning opportunities must be updated frequently and automatically. An aggregator system provides up-to-date information due to its automatic data harvesting functionality.

3.1 Target Audience

There are two kinds of key users of an aggregator portal which is based on MLO-AD and provides up-to-date information about learning opportunities: *Prospective learners* compare different programs and courses, while *providers of learning opportunities* advertise their offers.

3.1.1 Prospective Learners

The primary target audience of the portal based on the MLO-AD standard is comprised of learners who want to search, browse, and compare information about electronically represented learning opportunities. People who are interested in learning opportunities use the aggregator portal as a search engine to receive this special information.

The MLO-AD standard aims to address all types of learners: Children and their parents search for learning opportunities to achieve special skills or to have another leisure activity; pupils and students are looking for courses and study programs for their education; professionals want to expand their knowledge with further training; companies offer conferences or training sessions for their employees; or people just want to do courses in topics they are interested in. The MLO-AD defines various properties to differ entiate between the types of learning opportunities, which are described in the following way [39, pp. 9–11]:

- *Qualification* defines the qualification which can be obtained from completion of a learning opportunity.
- *Credit* includes an account of credits than can be obtained from completing a learning opportunity.
- *Level* indicates the intended outcome of the learning opportunity in terms of progression.
- *Engagement* describes how individuals engage in a learning opportunity. It encompasses temporal, modal and spatial patterns of engagement and attendance (e.g., overall attendance includes full-time and part-time; modes of study, which can be distance, campus-based, workplace-based, or online; pattern of attendance hours, like evenings, daytime, weekend).
- *Objective* is the aim or learning objective for the learning opportunity.
- *Prerequisite* describes the entry requirement for accessing the learning opportunity.

The aggregator portal based on the MLO-AD standard is tailored at first instance to professional learners and postgraduates who aim to expand their knowledge with professional courses. It can be assumed that the target audience has good computer or media skills, and is familiar with the use of search engines. The required information about learning opportunities and adequate user interfaces for all kinds of prospective learners should be provided in future development.

An enrollment in a course or any other educational program is mostly dependent on several criteria, which could be subject, certification or qualification, location, starting date, duration, costs, language, and prerequisites. Prospective learners want to get all the relevant information about learning opportunities for a subject of their interests. Information or content about educational programs that can be found in the Internet is often cluttered, which increases the need for organizing and accessing the information in a way that is useful, educational, and structured. Learners usually try to avoid visiting various websites to receive and collect the information they need. To get a good survey of opportunities and to be able to compare them, learners use the aggregator portal based on the MLO-AD standard. Naturally, this portal provides up-to-date information about educational offers. Prospective learners enter the information they are interested in into a search mask and receive the data of all available learning opportunities arranged by the data model of MLO-AD.

By providing comprehensive data of various learning opportunities, the aggregator portal also supports transparency for learners.

3.1.2 Learning Opportunity Providers

The second group of key users consists of providers of learning opportunities that include universities, colleges, training organizations, as well as experts who can offer professional courses. These providers wish to advertise their educational offers widely and in an economical way. With a portal based on the MLO-AD standard, providers of learning opportunities can reach their target audience across multiple countries.

Again, MLO-AD is a standard for all types of educational programs and, therefore, addresses a wide range of learning opportunity providers. The first version of the aggregator portal based on MLO-AD is developed for professional courses or training sessions, and accordingly for providers of these types of learning opportunities.

Currently, many aggregators that collect information about learning opportunities do not offer international standard formats. This means providers usually have to input their data manually on web forms or upload them with

specialized tools. Due to the fact that they are not based on a standard, existing aggregators sometimes require specialized vocabularies or encoding schemes (e.g., for identifiers) [39, p. 12]. To be able to provide their data for different services and, therefore, to advertise their educational offers, learning opportunity providers need to adjust their data model to the requirements of the existing service or have to provide the information "manually." Providing data about learning opportunities can become expensive for providers (especially if they want to supply their data for multiple systems) if they cannot base their data model on a widely accepted standard. As a result, learning opportunity providers will support services based on the MLO-AD standard to advertise their educational offers.

The aggregator portal can collect information about learning opportunities automatically from a provider if its data model is based on a standard like MLO-AD. The provider does not have to cope with updating its data about the educational offers on the portal itself. It just has to revise the information in its own database as the portal collects and updates this data frequently and automatically. This procedure ensures the consistency of the information about learning opportunities on all systems collecting this data from the provider.

Although the data is collected and updated automatically, providers usually still wish to have influence on their information which is available on the aggregator system. For example, they intend to update data or add new offers to the database of the aggregator immediately, and do not want to wait for the next update from the system. It is also possible that new providers want to advertise their learning opportunities using the aggregator portal based on MLO-AD, and need a way to indicate themselves. Therefore, the aggregator portal still needs to provide the possibility to input data about learning opportunities manually.

If an educational institution provides its information for the aggregator portal, learners can search for and browse its courses and programs and become aware of the institution. As the portal can be used across multiple countries, providers can also reach learners who are out of range through their usual advertisement of learning opportunities.

Learning opportunity providers want to extend the range of their advertisements not only by covering more countries, but also by using new and popular technologies. In March 2009, The Nielsen Company published a report [36] dealing with the internet consumer phenomenon *Social Networking* which facilitates the building of online communities of people who share the same interests. This internet activity is in such great demand that "two-thirds of the world's Internet population visit a social network or blogging site and the sector now accounts for almost 10% of all internet time" [36, p. 1]. Since

December 2008, social networking has had an even higher active reach than e-mail [36, p. 2], which can be attributed to the fact that many of these services provide e-mail or instant messaging to communicate with other users. Social network services like Facebook[1], MySpace[2], Hi5[3], LinkedIn[4], Orkut[5], Xing[6], and 51.com[7] are counted among the most popular websites[8] [9].

Obviously, advertisers are interested in the social networking market to appeal to new consumers. However, according to [36, pp. 5–6], "members have a greater sense of 'ownership' around the personal content they provide and are less inclined to accept advertising around it. [...] Advertising shouldn't be about interrupting or invading the social network experience, it should be part of this conversation."

Thus, learning opportunity providers can advertise their educational offers via social network services wisely if the user gets an added value by interaction. This can be provided not only through searching for a learning opportunity via an integrated gadget of the aggregator portal, but also by sharing educational offers with friends. If a user is interested in a specific training session, his friends may also like to attend this event.

Additionally, users can build a miniature social network according to a learning opportunity. They can exchange expectations, start discussions about the subject of the training program or give feedback to other users and the educational institution or instructor. The trainer or teacher of the learning opportunity can communicate directly with learners before or after a session, which is also a new way of advertising an educational event.

Educational institutions and instructors can advertise their learning opportunities in a target-oriented way with a gadget that provides the functionality of the aggregator portal based on MLO-AD and is integrated into a social network service. Gadgets are miniature web applications embedded into another website. As the first version of the aggregator portal has been developed for professional courses, social networks mainly used by professionals and postgraduates (like LinkedIn or Xing) are well applicable for this tool.

[1] http://www.facebook.com
[2] http://www.myspace.com
[3] http://www.hi5.com
[4] http://www.linkedin.com
[5] http://www.orkut.com
[6] http://www.xing.com
[7] http://www.51.com
[8] http://www.alexa.com/site/ds/top_sites
[9] http://mostpopularwebsites.net

3.2 Product

Currently, none of the existing aggregator portals providing information about learning opportunities are based on the MLO-AD standard. This is due to the fact that MLO-AD is still young and in the process of being standardized. Together with GTN-Québec[10], the idea of an aggregator portal based on MLO-AD was created to demonstrate the potential and practicability of this new standard and to fulfill the need of international comparison and transparency of learning opportunities.

The aggregator portal must meet certain requirements to come up to the expectations of the target audience. One main condition is the permanent accessibility of the information about learning opportunities all over the place. To provide this permanent access and to avoid data inconsistency (which can occur if the system cannot harvest the required data frequently), the aggregator portal is developed as a web application.

Further requirements are described in this section more precisely. They are separated into *system logic* (back-end) and *user interface* (front-end), according to their relation to the system.

3.2.1 System Logic

The back-end of a web system includes those components which process the output from a user interaction. It covers the logic of a system which happens on the server and mostly includes operations on the database as well. These procedures are usually hidden from the user.

The database model of the aggregator portal is based on the MLO-AD standard for describing learning opportunities. This ensures a wide acceptance of the system, an easy collection of data from providers supporting this standard, as well as the possibility to compare all provided learning opportunities.

One of the main components responsible for a logical process in the aggregator portal is the collection or harvesting of data from learning opportunity providers. Educational institutions must provide their data using a data model based on MLO-AD, so that the portal can collect this information. The web system knows the providers which the information about learning opportunities is harvested from. References to the sources of the collected data are saved in a list with the URL of the provider as a unique identifier. It also needs the permissions from the providers to be able to access the data of the learning opportunities. Different possibilities of harvesting information from an educational institution are described in Chapter 4. Considering the

[10] http://www.gtn-quebec.org

complexity of MLO-AD (multiple elements with different data types), more questions about data harvesting arise:

- Which elements of the standard are collected and must therefore be provided by the educational institution?
- Is there a need to provide vocabularies for certain elements to make them comparable with those of other learning opportunities?
- How can it be ensured that the providers comply with these vocabularies?
- How is it possible to avoid spam and to ensure a good quality of the information to describe learning opportunities?

Section 4.7 recommends a solution to harvest information for the aggregator portal based on the MLO-AD standard and considers the mentioned questions.

Besides automatically collecting the data, the aggregator portal provides an easy and quick way for manually inputting information about learning opportunities. This solution is described in Section 4.1.

To keep the data of the aggregator portal consistent with the information provided by the educational institution, the portal harvests this information at regular intervals. The calculation of the frequency of these updates is another difficulty that has to be solved by the system. Providers of learning opportunities usually adjust the data according to the types of their offers. There is no need to update and collect the data every week from an educational institution which changes its offers twice a year. However, it is always possible that the information of learning opportunities changes at any time or that new training programs are offered unexpectedly. It needs to know if the information of a learning opportunity has changed to avoid the unnecessary updating of data. A solution for collecting the information at adequate time intervals and for avoiding unnecessary data transfer is described in Chapter 4.

While the aggregator portal updates existing data about learning opportunities, it also considers new educational offers. Therefore, the system must be able to identify new learning opportunities. It also detects and deletes old and expired information about courses or training programs. As mentioned before, the provided content of the aggregator portal based on the MLO-AD standard must be comprehensive, up-to-date, and without spam or expired information.

Another logical part of the aggregator portal is the support of targeted searches. The system searches its database for learning opportunities ac-

cording to the user's data input. Thereby, it allows the search by various keywords and the sorting of results by different criteria. This search procedure must operate efficiently to be able to quickly provide the user with the required information.

3.2.2 User Interface

The front-end of a web system manages the user interaction and forwards its data to the components of the back-end. It shows the user interface and displays results of an interaction given by the logic of the system.

The user interface of the aggregator portal for learning opportunities must allow prospective learners to search and browse for educational offers. The website provides a simple and quick way to start the search with two input fields: One is for keywords about the subject of the learning opportunity and the other one for the desired location. The optional input field for the location is a drop-down menu which allows multiple selections. Additionally, users can extend their input mask to provide further information they are interested in and to limit the search results. Examples for data to refine the search are: name of the educational institution, starting date, duration, costs, language of instruction, engagement, prerequisites, and qualification. Advanced users can directly input various information about a learning opportunity into the main subject field by using defined keywords with the desired values as a query. Hence, they can refine their search without extending the mask of the input fields for further information. The main user interface of the search mask is simple, but functional for the refined search. It is similar to Google, which is extendable by Google Advanced Search[11]. As the target audience of the prototype only includes professionals and postgraduates, the type of program is not considered in the search user interface for the first version of the aggregator portal.

After the search process has been completed, the user interface shows the comprehensive information about the detected learning opportunities and their providers in an appropriate way. Thereby, users have the possibility to change the view according to the information they are especially interested in. Usually, the displayed results are grouped by the institutions, which are the providers of the learning opportunities. However, if prospective learners want to attend an educational program just within a special location, they can switch the view to see the results displayed with flags on a map, as in Google Map. Another possibility for an alternative view is to show the results within a calendar according to the starting date and duration of the learning opportunity. The calendar view additionally provides users with a

[11] http://www.google.com/advanced_search

good survey about whether an educational institution offers multiple sessions or dates for a training program.

Moreover, users are able to sort the results according to the information of their special interests. This can be a sort by program type, costs, starting date, duration, or any other criterion supported by MLO-AD. Users can adjust the presentation of the search results to their needs. These possibilities facilitate the transparency and comparison of learning opportunities for each learner.

The website of the aggregator portal based on MLO-AD also supports a browsing of learning opportunities. The browser user interface divides the educational offers according to their subjects and categories. Moreover, the system allows a pre-selection of the location or country. Within a category, the user can also fetch a specific number of upcoming learning opportunities with the selected subject.

The aggregator portal based on MLO-AD is additionally available as a gadget for social network services. This ensures a higher and targeted reach of prospective learners, because of the integration into popular social network services that are used by the target audience. The gadget is a miniature of the aggregator portal, and the user interface of the search mask is the same. If a user searches for an educational event via the gadget, he/she will receive only a limited number of results. More, resp., all results are accessible via a link to the "normal" user interface of the aggregator portal. The gadget does not support the browsing of learning opportunities.

However, it provides other services matched to users' behaviors in social networks. The aggregator portal integrated in a social network allows users to share learning opportunities with their friends or contacts. They can invite or encourage each other to participate in a specific educational event. A click on the button "Share with Contacts" near the learning opportunity data opens the list of all the users' contacts. Accordingly, users can choose which contacts they want to share the specific training with. Users can also add a session of a learning opportunity to their "activities". Consequently, all their contacts can note that they have been attending these events. Especially in social networks such as LinkedIn and Xing, which are mainly used by professionals, users specify their abilities, education, and training programs they participated in. Information about learning opportunities from the aggregator portal based on MLO-AD can be used as a reference.

Groups connected with a learning opportunity can also be created or participated in. This feature gives users the possibility to discuss an event, post comments, or ask questions to the instructor (who is also a user of the social network).

Beside the solutions to difficulties relating to the user interface, Chapter 5 of this work describes the integration of the gadget into social network services.

3.3 Success Criteria for the Prototype

The first step of developing the aggregator portal based on the MLO-AD standard is the creation of a prototype. Chapter 6 of this paper includes the analysis of the prototype development process and Chapter 7 evaluates the results according to the requirements.

The prototype must comply with the following criteria:

- The database and its data model for describing learning opportunities are based on the MLO-AD standard.
- Data about educational offers are harvested from at least two learning opportunity providers.
- One of these providers is the training center of CRIM[12].
- The portal aggregates the information automatically and in adequate time intervals. In this process, it not only updates its existing data, but also covers information about new, recently emerging educational offers.
- The portal allows targeted searches for learning opportunities.
- The prototype is implemented as a gadget for social network services.
- The gadget provides a user interface which allows the user to enter the information about a learning opportunity he/she wants to search for. The search mask consists of the main input field for keywords about the subject.
- The results of a search for learning opportunities are shown in a simple way. The possibilities to change the view of the given results or to sort the detected data by special criteria are not priorities for the prototype.

The aggregator portal based on the MLO-AD standard is developed as an open source project in order to encourage the further development of the prototype and to increase the popularity of the system. The project is licensed under the *Educational Community License, Version 1.0* of Open Source Initiative (OSI), who define themselves on their website[13] as "the stewards

[12] http://www.crim.ca/en/services/Formation
[13] http://www.opensource.org/about

of the Open Source Definition (OSD) and the community-recognized body for reviewing and approving licenses as OSD-conformant." The terms and conditions of the Educational Community License[14] include:

> Permission to use, copy, modify, merge, publish, distribute, and sublicense this Original Work and its documentation, with or without modification, for any purpose, and without fee or royalty to the copyright holder(s) is hereby granted, provided that [...] the following [see full terms and conditions of License for following; author's note] is included on ALL copies of the Original Work or portions thereof, including modifications or derivatives, that [are made].

[14] http://www.opensource.org/licenses/ecl1.php

Chapter 4

Data Collection from Learning Opportunity Providers

The core functionality of the aggregator portal based on MLO-AD is the collection of its data to describe learning opportunities. Data collection is carried out automatically and at regular intervals to keep the repository of the aggregator portal up-to-date. Consequently, the learning opportunity provider does not have to cope with managing its information provided on the aggregator portal.

This chapter explains different technologies to collect information from data repositories. The core functionalities of these technologies are described to gain a better insight into which circumstances they are applicable for data specified by MLO-AD. Each section also includes an analysis of the usability of the specific technology for MLO-AD.

In the future, the aggregator portal will support different ways of collecting data about learning opportunities to address more educational institutions and facilitate the provision of their information.

4.1 Manual Provision of Data

Usually, the aggregation of information about learning opportunities and their providers is carried out automatically and regularly via a data collection technology. However, special cases can occur, so that learning opportunity providers wish or need to deliver their information manually.

One of these cases could be an unexpected update or creation of a learning opportunity, whereas the educational institution wants to submit these changes immediately to the aggregator system without waiting for the next

automatic update by the system. Another possibility could be that a new learning opportunity provider wants to advertise its offers using the aggregator portal based on MLO-AD and needs a way to indicate itself. And, of course, errors which cannot be detected or solved by the system (e.g., missing or conflicting data, wrong mapping of information about a learning opportunity object to an MLO-AD property) can occur during the data harvesting process. In this case, the learning opportunity provider or the administrator of the aggregator portal has to update the information manually.

The most common way for manually providing the data on the Web is via forms with input fields. The aggregator portal provides a form on a website with an input field for each property of a learning opportunity, as well as of the educational institution. The form builds the framework of the MLO-AD data model and ensures a proper mapping of information to the properties. Additionally, it allows default values, specific vocabulary, or particular data types for certain properties to be defined. However, web forms are often misused to clog data repositories with spam or to submit wrong information. To retain the quality of the information provided by the aggregator portal based on MLO-AD, it needs an identification of the data provider. This means an educational institution has to register and log-in before submitting data about learning opportunities via the form.

Another problem relating to the usability of the web form can occur. If an educational institution has to provide its information manually, it usually wishes to do this in an easy and quick way. By using the web form, the learning opportunity provider has to log-in, input all information into the specific field of the property, and submit the form for each update. This is probably an adequate effort to modify properties about one learning opportunity. However, if the institution wants to update more educational events, it will wish to use another possibility for the manual provision of data, especially if its offers consist of similar data.

This alternative possibility of manual data provision is a technique via e-mail similar to the Mail-to-Blogger system of Blogger[1]. Blogger users can post a blog entry by sending an e-mail to a certain e-mail address. Of course, an update or creation of a learning opportunity specified by MLO-AD is more complex than a post of a blog entry. Therefore, the learning opportunity provider needs to use a template, which consists of the MLO-AD properties, in the e-mail. This template will be available for download on the web page of the aggregator portal. The educational institution adds the specific values of the learning opportunity to the properties of the template. Obviously, this technique is more error-prone, compared to the web form. The aggregator portal receiving and processing the e-mail must be able to handle errors, which could be not using a template or making mistakes in the template.

[1] http://help.blogger.com/bin/answer.py?answer=41452

However, the technique by means of e-mail can be more convenient for the data provider, especially when more updates must be submitted. Data in the MLO-AD template can be easily copied and modified for another update. To ensure the quality of the information provided by the aggregator portal, as well as to identify the provider of this information, the educational institution has to register once on the system. Each institution gets a unique e-mail address, which can be used to submit data by e-mail. An e-mail address for each learning opportunity provider allows an easy identification of the e-mail sender, which, in turn, is also a security check from the system.

Both techniques of manual data provision have to face the difficulty of detection if the input data is an update of a learning opportunity object or the creation of a new one. Therefore, each learning opportunity needs a unique identifier which meets the standardized form of a Uniform Resource Identifier (URI). This URI includes the path of the data source that also identifies the learning opportunity provider. Hence, this information can additionally be used to check the quality of the supported data. The aggregator portal has to check the URI of a learning opportunity included in the manual data provision to see if it is already available in the system's data repository. If so, it will overwrite the available data with the submitted information or otherwise create a new resource of a learning opportunity.

4.2 Web Feeds

A *web feed*, *news feed* or simple *feed* is a document often based on XML (Extensible Markup Language) and used to transfer frequently updated content to users [42, p. 78]. This content can vary from weblogs entries, or information about video and audio files (these feeds are called "podcasts"), to news items or any kind of content that can be packaged into a unit. Feeds have gained in popularity with the increased use of weblogs.

Many aggregators use feeds to collect data from multiple sources (see Section 2.4). Before an aggregator can harvest information via web feeds, a content provider has to produce a feed and publish the feed link on its website. The feed is updated every time the content changes. The aggregator (also called "feed reader") subscribes to the feed by saving this link in its feed list next to links of other feed providers. At scheduled intervals or when instructed, the aggregator iterates this list and asks for new content. If new content is available, the feed reader downloads this information by using the links to the sources saved in the web feed.

Web feeds are processed through pull technology, which means that the user (or his/her aggregator) is responsible for receiving new content from a provider. The most popular XML news feed formats are *RSS* and *Atom*.

4.2.1 RSS

RSS was first released in 1999 and had developed during the following three years into two main branches: RSS 1.0 is based on the Resource Description Framework (RDF), which is part of the Semantic Web, while RSS 2.0 is the current heir to the line of XML formats [42, p. 78]. This section deals with RSS 2.0[2], whereas RSS is the abbreviation for *Really Simple Syndication*.

The basic element of the RSS XML document is the `rss version="2.0"` element which is followed by a single `channel` element. The latter represents the source of the feed and contains the entire feed content as well as all associated metadata [17, Sec. 4.2]. A `channel` is described by three required elements, which are `title`, `link`, and `description`, and can include any number of `item` elements [42, p. 79]. Optional elements of `channel` can be, for example, `language`, `copyright`, `category`, `pubDate` (which is the publication date of the content), `lastBuildDate` (the date and time when any item of the RSS feed was last changed), and `image` (which has sub-elements including `url`, `title`, and `link`). An `item` element contains the primary content of the feed [17, Sec. 4.2] and must include at least a `title` or a `description`. Other sub-elements of an item are, amongst others, `link`, `author`, `category`, `pubDate` and `source` (from which the item was derived).

The following XML file is a simple RSS feed describing two learning opportunities (whereas it does not use the specifications of MLO-AD). Entity-encoded HTML is used in the `description` elements.

```
1 <?xml version="1.0"?>
2 <rss version="2.0">
3   <channel>
4     <title>CRIM Training Center</title>
5     <link>http://www.crim.ca/en/services/Formation</link>
6     <description>The CRIM Training Center provides a highly effective
          learning environment that allows learners to benefit from
          superior quality training.</description>
7     <language>en-us</language>
8     <pubDate>Mon, 2 Mar 2009 10:12:12 EST</pubDate>
9     <lastBuildDate>Mon, 2 Mar 2009 10:12:12 EST</lastBuildDate>
10    <item>
11      <title>Object-Oriented programming with C#</title>
12      <link>http://www.crim.ca/en/services/Formation/Cours-inscription/
            index.html?uri=/en/services/Formation/Cours-inscription/
            recherche.html&id=NET513en</link>
13      <description>&lt;p&gt;&lt;i&gt;C#&lt;/i&gt; is the most popular
            language in &lt;i&gt;Microsoft .NET&lt;/i&gt;. It is a pure
            object oriented (OO) language: to develop with C#, we must
            think in terms of objects. Global variables and global
            functions don't exist: everything is a class. Therefore, to
```

[2] http://cyber.law.harvard.edu/rss/rss.html

```
            develop in C# we must use OO concepts and know how to apply
            them effectively.&lt;/p&gt;</description>
14      <pubDate>Mon, 2 Mar 2009 09:05:15 EST</pubDate>
15    </item>
16    <item>
17      <title>Overview ISO 20000</title>
18      <link>http://www.crim.ca/en/services/Formation/Cours-inscription/
            index.html?uri=/en/services/Formation/Cours-inscription/
            recherche.html&id=ITI611en</link>
19      <description>&lt;p&gt;This presentation offers a good overview of
            the only international standard that offers official
            recognition for a business in IT service management.&lt;/p&gt
            ;</description>
20      <pubDate>Mon, 2 Mar 2009 10:06:01 EST</pubDate>
21    </item>
22  </channel>
23 </rss>
```

Like XML, RSS is also extensible. This means that new elements which are not declared by the RSS specification can be added. The only condition for extending an RSS is to define the added elements in a namespace. This possibility makes RSS files more flexible, but also more difficult to parse and read [21, pp. 67–68].

4.2.2 Atom

The *Atom Publishing Format* (simply called Atom) was released as an Internet standard in 2005 by the Internet Engineering Task Force (IETF)[3] [21, p. 70]. The current version is specified as Atom 1.0[4].

Although the semantics of Atom 1.0 are similar to RSS 2.0, they have a different naming scheme [42, p. 83]. The following Atom document describes the same content as the previous RSS 2.0 file.

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <feed xmlns="http://www.w3.org/2005/Atom">
3   <title>CRIM Training Center</title>
4   <subtitle>The CRIM Training Center provides a highly effective
        learning environment that allows learners to benefit from superior
         quality training.</subtitle>
5   <link rel="alternate" type="text/html" href="http://www.crim.ca/en/
        services/Formation"/>
6   <link rel="self" href="http://www.crim.ca/en/services/Formation/Atom1
        .0_Example.xml"/>
7   <updated>2009-03-02T10:12:12-05:00</updated>
8   <author>
```

[3] http://www.ietf.org
[4] http://www.ietf.org/rfc/rfc4287.txt

```
 9     <name>CRIM</name>
10     <email>info@crim.ca</email>
11   </author>
12   <id>tag:www.crim.ca,2009:/en/services/Formation</id>
13   <entry>
14     <title>Object-Oriented programming with C#</title>
15     <link href="http://www.crim.ca/en/services/Formation/Cours-
           inscription/index.html?uri=/en/services/Formation/Cours-
           inscription/recherche.html&id=NET513en"/>
16     <id>tag:www.crim.ca,2009:/en/services/Formation/Cours-inscription/
           recherche.html&id=NET513en</id>
17     <updated>2009-03-02T09:05:15-05:00</updated>
18     <summary type="html">&lt;p&gt;&lt;i&gt;C#&lt;/i&gt; is the most
           popular language in &lt;i&gt;Microsoft .NET&lt;/i&gt;. It is a
           pure object oriented (OO) language: to develop with C#, we must
           think in terms of objects. Global variables and global functions
            don't exist: everything is a class. Therefore, to develop in C#
            we must use OO concepts and know how to apply them effectively
           .&lt;/p&gt;</summary>
19   </entry>
20   <entry>
21     <title>Overview ISO 20000</title>
22     <link href="http://www.crim.ca/en/services/Formation/Cours-
           inscription/index.html?uri=/en/services/Formation/Cours-
           inscription/recherche.html&id=ITI611en"/>
23     <id>tag:www.crim.ca,2009:/en/services/Formation/Cours-inscription/
           recherche.html&id=ITI611en</id>
24     <updated>2009-03-02T10:06:01-05:00</updated>
25     <summary type="html">&lt;p&gt;This presentation offers a good
           overview of the only international standard that offers official
            recognition for a business in IT service management.&lt;/p&gt
           ;</summary>
26   </entry>
27 </feed>
```

Atom's basic element is `feed`, which uses an Atom-related XML namespace[5] and contains one or more `entry` elements as items. An Atom feed must include a `title`, the `updated` element (specifies the date of the last modification), and the `author` [21, p. 71]. `<link rel="alternate" ... />` indicates that this document is a feed of the declared website, while `<link rel="self" ... />` specifies the link of this feed document [42, p. 84]. These links are also mandatory for the `feed` element. Entries are described by tags of `title`, `link`, `id`, and `summary`, whereas the first three elements are mandatory. Whereas RSS uses descriptions to describe the content of an item, Atom offers two elements: `summary` and `content`. Especially if the content is non-textual or non-local (e.g., identified by a link), the summary is important for accessibility reasons. An entry must not have more than one `summary` and `content` element. Like the `feed` element, `entry` must also include an

[5] http://www.w3.org/2005/Atom

`updated` element. The `type` attribute in the `summary` element of an entry indicates that this summary includes entity-encoded HTML. Other types are "text" (no HTML included), "xhtml," "application/rdf+xml," and external type specified by "application/xyz" and its source. This specification of content types is called "Atom content model," and is even more important in the `content` element.

Atom is extensible in the same way as RSS 2.0, which means that elements can be included if they are defined in an XML namespace. Compared to RSS, Atom is more specified, and it is likely to replace RSS in the future, since the specifications of RSS will not be further developed or clarified [21, p. 78].

4.2.3 Web Feeds for MLO-AD

As RSS and Atom feeds can include any content, they can also transfer updated information about learning opportunities. The possibility to extend these XML documents allows more elements that can specify properties of a learning opportunity or its provider to be added.

To make web feeds applicable for MLO-AD, it needs to match the required properties of MLO-AD objects to existing RSS or Atom tags. Due to the reason that both feed formats do not offer enough suitable elements to match all properties, new elements that use an XML namespace must be defined additionally. Fig. 2.1 in Section 2.2 shows the complexity of the MLO-AD model. Learning opportunity providers offer learning opportunity specifications that are an abstract description and specify one or more learning opportunity instances. This shows that the MLO-AD resources stand in relation to each other. Moreover, each object specifies various properties.

Web feeds are XML documents with simple specifications. Atom is the technically superior feed format and better qualified for sophisticated requirements like those of MLO-AD [21, p. 179]. Therefore, it is recommended to use Atom feeds for collecting data specified by MLO-AD. However, it is also a challenge to realize an efficient mapping of the relation between a learning specification and an instance with Atom. Due to the reason that feeds are generated by the providers (which are various learning opportunity providers in this case), there is no possibility by web feeds to ensure that all feeds of the different providers use the same structure for describing learning opportunities based on MLO-AD. Therefore, parsing and "understanding" these feeds correctly is a challenge for the aggregator portal based on MLO-AD.

However, web feeds are commonly used to transfer updated content on the Internet. Especially Atom can be considered as a data collection technology for MLO-AD, even if the aggregator has to face some difficulties when parsing these special documents.

4.3 Semantic Web

> The *Semantic Web* is not a separate Web, but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation,

says Tim Berners-Lee (the inventor of the World Wide Web), James Hendler, and Ora Lassila in the article "The Semantic Web," published in *Scientific American* [7].

The current Web is built on HTML and XML, which describes the structure of the information presented for humans on a website, and does not include any information about the meaning of the displayed data [2, pp. 37–38]. Today's search engines index HTML pages to find answers, but they also return a lot of irrelevant information. They just look at occurrences of words in documents which is a hint, but does not tell what the document really is about [6, pp. 177–178]. The meaning of the web content is not automatically processable or "understandable" by computers or applications. "Smart" web applications or software agents that solve complex problems can only be as intelligent as the data that is available to them [3, p. 3].

The Semantic Web provides a knowledge representation of linked data in order to allow machine processing. This representation of knowledge is realized through information about information (which is also called "metadata"), as well as through connections between different forms of data [6, pp. 181, 185]. Web applications and agents can use all kind of data on the Web by connections and by using rules to conduct automated reasoning. This added logic of the Semantic Web must enable the description of complex properties of objects, but must not be too intricate that machines or agents can be tricked by being asked to consider a paradox [2, p. 69]. The Semantic Web and its knowledge representation of linked data is more practical than the current Web because applications can get the data they need [3, p. 4]. Tim Berners-Lee's vision of the Semantic Web is as described in [7]:

> The real power of the Semantic Web will be realized when people create many programs that collect web content from diverse sources, process the information and exchange the results with other programs. The effectiveness of such software agents will increase exponentially as more machine-readable web content and automated services (including other agents) become available. The Semantic Web promotes this synergy: even agents that were not expressly designed to work together can transfer data among themselves when the data come with semantics.

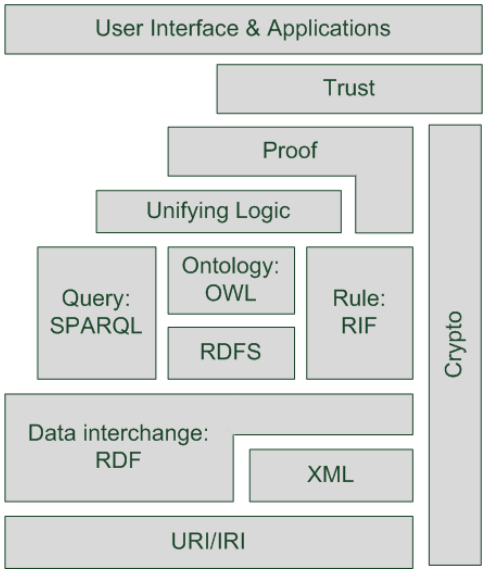


Figure 4.1: Semantic Web Layer Cake (according to Tim Berners-Lee's diagram published on the website of W3C Semantic Web Activity)

To be able to add logic to the Web, the Semantic Web needs different technologies and standards which are organized in layers built one upon another. The architecture of the Semantic Web is illustrated by the *Semantic Web Layer Cake* or *Semantic Web Stack*, which is shown in Fig. 4.1. Building one layer upon another requires each layer to be aware of a layer to interpret at a lower level and to take at least partial advantage of information at higher levels [2, p. 69].

The Semantic Web is based on Uniform Resource Identifiers (URI) and Internationalized Resource Identifiers (IRI) to define web resources (all things on the Web that can be identified) in a unique way. XML (Extensible Markup Language) is responsible for structuring the data of a resource, is extensible by arbitrary tags, but says nothing about the meaning of the information or structure. XML is an open standard to exchange data between applications over the Web, similar to HTML, which allows information to be displayed and exchanged over the Internet. It is also the bridge to exchange data between the two main web software development frameworks: J2EE and .NET [2, p. 69]. XML is often combined with an XML schema which defines the structure of the XML document and extends it with data types. The Resource Description Framework (RDF) is the basic framework of the Semantic Web [3, p. 28] and expresses data models which refer to resources and their relationships in a formal way to enable software agents to read and process them. RDFS (or RDF-S) is the abbreviation for RDF Schema, which provides basic elements for the description of RDF vocabularies and

structures RDF resources. SPARQL stands for SPARQL Protocol And RDF Query Language, and is a World Wide Web Consortium (W3C)[6] recommendation that has been used since January 2008 to query data based on RDF [29]. The Web Ontology Language (OWL) belongs to the knowledge representation language family for authoring ontologies in the Web. Ontologies are collections of information and define relations among terms [2, p. 55]. Rule Interchange Format (RIF) is still being developed[7] by a W3C working group, and will bring rules support to the Semantic Web. The upper layers of the Semantic Web that are underneath the user interface (which is the connection to the user) are trust, proof, and logic, which should prevent wrong information or relations as well as spam pages or spam ontologies. These layers have not been fully realized yet. An organization of documents by chains of trust will support the identification of trustful information.

4.3.1 RDF

The *Resource Description Framework* (RDF) is a standard developed by the W3C for representing information about resources in the World Wide Web [25, Ch. 1].

Each resource is identified by a URI. RDF enables statements to be represented in the form of subject-predicate-object sentences, which are also called RDF *triples*. A triple relates a subject to an object via a predicate, while all three elements are identified by URIs [35, p. 83]. Fig. 4.2 shows an example of a triple, whereas "Learning Opportunity Provider" is the subject described by the statement, "Offers" is the property of the subject, and "Learning Opportunity" is the value of the statement. This simple model of the triple used by RDF has many advantages. One of the most important is described in [2, pp. 87–88] in the following way:

> Any data model can be reduced to a common storage format based on a triple. This makes RDF ideal for aggregating disparate data models, because all the data from all models can be treated the same. This means that information can be combined from many sources and processed as if it came from a single source.

Multiple triples can be connected to form an RDF *graph*, whose nodes illustrate URIs of resources and whose arcs are properties [35, p. 84]. Triples of a graph can also originate from different sources. This supports the idea of the AAA slogan that says: "Anyone can say Anything about Any topic" [3, p. 35], meaning that anyone can create a statement about any resource.

[6] http://www.w3.org
[7] http://iit-iti.nrc-cnrc.gc.ca/new-neuf/2008/08-09-09_e.html

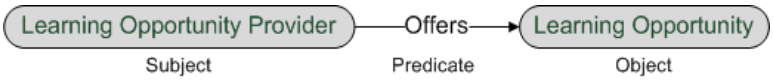


Figure 4.2: Example of an RDF triple

RDF provides an XML-based syntax (RDF/XML) for recording and exchanging triples or graphs [25, Ch. 1]. The following example shows the triple of Fig. 4.2 in an RDF/XML syntax:

```
1 <?xml version="1.0"?>
2 <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3          xmlns:mlo="http://www.example.org/mlo/">
4   <rdf:Description rdf:about="http://www.example.org/
        learningOpportunityProvider">
5     <mlo:offers>Learning Opportunity</mlo:offers>
6   </rdf:Description>
7 </rdf:RDF>
```

The `rdf:RDF` element in line 2 (and which ends in line 7) declares its content as RDF. XML namespace declarations indicate that all tags prefixed with `rdf:` are part of the namespace identified by the URI http://www.w3.org/1999/02/22-rdf-syntax-ns# [25, Sec. 3.1], and tags using the prefix `mlo:` are MLO-AD elements. Lines 4–6 provide the RDF/XML for the specific statement, which is represented as a `Description` that is `about` a subject (in this case, about http://www.example.org/learningOpportunityProvider). The content of `rdf:Description` elements are called property elements and may contain other descriptions producing nested descriptions [2, p. 91]. This example includes just one property, which is `mlo:offers` with the value `Learning Opportunity` described as a plain literal. An RDF/XML file can cover statements about multiple subjects, whereas each subject is described by an `rdf:Description` element.

An RDF document can refer to an RDF schema (RDFS), which provides the facilities to define vocabularies and to indicate specific classes and properties of resources, as well as the way they must be used together [9]. Consequently, the structure of the RDF document is specified by the RDFS.

RDF data can be queried and accessed via SPARQL. This referenced standard consists of three specifications: a description of the language to query RDF data across diverse data sources [29], a protocol which is described with WSDL 2.0 as well as by HTTP and SOAP bindings to query remote databases [11], and the XML format of the query result which will be returned [5].

4.3.2 OWL

The *Web Ontology Language* (OWL) for knowledge representation has been developed and standardized by the W3C and is described in [26] as follows:

> OWL is designed for use by applications that need to process the content of information instead of just presenting information to humans. [...] OWL can be used to explicitly represent the meaning of terms in vocabularies and the relationships between those terms. This representation of terms and their interrelationships is called an *ontology*. OWL has more facilities for expressing meaning and semantics than XML, RDF, and RDF-S, and thus OWL goes beyond these languages in its ability to represent machine interpretable content on the Web.

OWL provides three sub-languages for different levels of expressiveness and efficient reasoning: OWL Lite, OWL DL, and OWL Full. OWL Full is like an extension of RDF, while OWL Lite and OWL DL (Description Logics) are like extensions of a restricted RDF [26, Sec. 1.3]. OWL-DL is the most prominent language of the OWL family and is most supported by the Semantic Web community [35, p. 88].

OWL enhances RDF (which also provides some specifications for ontologies) inter alia with more vocabulary for describing properties, classes, as well as relations (e.g., disjointedness), cardinality, equality, and enumerated classes [2, p. 107]. The basis of OWL is formed by classes, relations between them, properties of classes, and constraints on relations and properties [2, p. 111].

The following code is a simple example of OWL in the RDF/XML syntax. It maps a part of the MLO-AD model shown in Fig. 2.1 of Section 2.2. The class "LoProvider" (Learning Opportunity Provider) offers a "LoSpecification" (Learning Opportunity Specification) class, whereby both classes are subclasses of class "LoObject" (Learning Opportunity Object).

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <rdf:RDF
3   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4   xmlns:rdfs="http://www.w3.org/200/01/rdf-schema#"
5   xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
6   xmlns:owl="http://www.w3.org/2002/07/owl#"
7   xmlns="http://www.example.com/mlo-ad.owl#">
8
9   <owl:Ontology rdf:about="http://www.example.org/mlo-ad">
10    <owl:versionInfo>Example of MLO-AD</owl:versionInfo>
11  </owl:Ontology>
12
13  <owl:Class rdf:ID="LoObject">
```

```
14    <rdfs:Label>Learning Opportunity Object</rdfs:Label>
15    <rdfs:comment>Abstract resource used for learning opportunities.</
          rdfs:comment>
16  </owl:Class>
17
18  <owl:Class rdf:ID="LoProvider">
19    <rdfs:Label>Learning Opportunity Provider</rdfs:Label>
20    <rdfs:subClassOf rdf:resource="#loObject"/>
21  </owl:Class>
22
23  <owl:Class rdf:ID="LoSpecification">
24    <rdfs:Label>Learning Opportunity Specification</rdfs:Label>
25    <rdfs:subClassOf rdf:resource="#loObject"/>
26  </owl:Class>
27
28  <owl:ObjectProperty rdf:ID="offers">
29    <rdfs:domain rdf:resource="#loProvider"/>
30    <rdfs:range rdf:resource="#loSpecification"/>
31  </owl:ObjectProperty>
32
33  <owl:DataTypeProperty rdf:ID="location">
34    <rdfs:domain rdf:resource="#loProvider"/>
35    <rdfs:range rdf:resource="&xsd;string"/>
36  </owl:DataTypeProperty>
37
38  <LoProvider rdf:ID="CRIM"/>
39 </rdf:RDF>
```

The header of the OWL document includes `rdf:RDF` as the root element and specifies a number of namespaces [2, p. 112]. The header also contains the `owl:Ontology` block for describing the current ontology. This block can include import statements by using `owl:imports` [26, Sec. 2.1]. Much of the power of ontologies comes from class-based reasoning, which is supported in OWL by the `owl:Class` elements [2, p. 113]. Classes are sets containing members that are also called individuals. Line 38 defines the individual "CRIM" for the class "LoProvider" specified in line 18. `rdfs:subClassOf` is a fundamental constructor for classes and relates a specific class to a more general one [22, p. 102]. These relations are shown in line 20 and 25. Properties are separated into datatype properties, which define relations with RDF literals or XML schema vocabularies (lines 33–36), and object properties, which relate instances of two classes (lines 28–31) [22, p. 103].

OWL provides many more constructs to specify classes, properties, their relations, and characteristics, as well as to define cardinality, equality, or restrictions. All constructs are precisely defined on the reference website[8].

[8] http://www.w3.org/TR/owl-ref

4.3.3 Semantic Web for MLO-AD

The Semantic Web provides different layers of technologies which can be used to describe and collect information about learning opportunities. Resources of MLO-AD can be described very well via RDF or OWL, since describing data models and their relations is the main task of these specifications. RDF and OWL also allow the inclusion of XML or RDF schemas that define the structure of a document to be able to describe a learning opportunity according to MLO-AD. Both specifications are based on XML, which is the standard to exchange data between applications.

Compared to OWL, RDF and RDFS have some disadvantages: It is not possible to define properties of properties, conditions for class membership, or equivalence and disjointness of classes [2, p. 102]. Additionally, "RDF is roughly limited to binary ground predicates and RDF Schema is roughly limited to a sub-class hierarchy and a property hierarchy with domain and range definitions," according to [2, p. 107]. However, OWL is a strong, but also complex language. RDF and RDFS are too simple for describing complex ontologies for the Semantic Web, but they are sufficient to describe the data model of MLO-AD.

To be able to collect information about learning opportunities via RDF, the provider of the educational offer has to create and publish this document on the Web. The aggregator checks this document for modifications at regular intervals. Unlike the Atom web feed, RDF or OWL do not specify an `updated` construct to identify a change of the document. To avoid the unnecessary transfer of data, it needs to specify an indicator (e.g., timestamp) to know if the document and its content were updated.

To be exact, RDF and OWL are both specifications to describe resources and are no standards for data transfer over the Web. However, RDF can be queried remotely by using the SPARQL protocol which is based on a web service technology. Another possibility to transfer RDF documents is via RDF feeds, which are the second development of RSS feeds specified as RSS 1.0 (RDF Site Summary)[9]. RDF feeds are not widely used, but examples do exist, as the website of Unified Data Feed[10] shows.

A drawback of RDF is the complex specification and, consequently, the complicated creation of RDF documents (OWL is even more complex than RDF). This could be a barrier for learning opportunity providers to describe their MLO-AD information by RDF.

[9] http://web.resource.org/rss/1.0/spec
[10] http://web2express.org/ufeed

4.4 Web Scraping

Web scraping is a technique to extract information from web pages, and is consequently also called *web data extraction*. It is often equalized with screen scraping and is closely related to web indexing. Most search engines use web robots or web spiders (which are a special type of a web robot) for web indexing and categorizing the content, which allows a faster provision of search results [41, pp. 104–110]. A web robot (also called "bot" or "crawler") is described in [19, p. 1] as "an Internet-aware program that can retrieve information from specific locations on the Internet." Compared to web indexing, web scraping is more focused on the transformation of unstructured content into structured data. Especially aggregators which collect information to offer price comparison or current weather statistics use web scraping to harvest their required data (see Section 2.4).

Before extracting data from a web page, a bot has to fetch the latest instance of the page from its respective URL and then parses the web page [8, p. 568]. According to the task of the bot, it looks for specific keywords and processes this data, transforms web content into structured data, or constructs and analyzes the HTML tree of a web page. This might seem to be a trivial process at first glance, but web scraping can also face some difficulties, as described in [28, p. 635]:

> A comprehensive data extraction process must deal with such obstacles as session identifiers, HTML forms, client-side JavaScript, incompatible datasets and vocabularies, and missing and conflicting data. Proper data extraction also requires solid data validation and error recovery to handle data extraction failures.

Several types of software for web scraping, which also support various features (e.g., transformation into XML, creation of web feeds, DOM parsing) to harvest data from web pages, already exist[11].

The process of web scraping, especially of complex data, can be facilitated for bots if the web page is enhanced with semantic information. Hereby, the web page is still based on HTML or XHTML (the "real" Semantic Web is based on RDF), but semantics are added by so-called "inline metadata formats." The two main technologies for this inline metadata are *RDFa* and *microformats*. Also *HTML 5*[12] provides more semantic information through new HTML elements, but it is not fully supported by all common browsers yet.

[11] http://en.wikipedia.org/wiki/Web-scraping_software_comparison
[12] http://www.w3.org/TR/html5

4.4.1 RDFa

RDFa is a W3C recommendation and allows human-readable data to be marked up with machine-readable indicators with a few simple XHTML attributes [1, Ch. 1]. Therefore, browsers and other agents can interpret and use this metadata. RDFa is only available for XHTML but not for HTML, since the latter is not extensible. The extension of XHTML results from reusing attributes from XHTML `meta` and `link` elements and applying them to other XHTML elements. Accordingly, this allows an annotation of XHTML markup with semantic information [15, p. 2]. RDFa uses and requires a new form of URIs called CURIE (compact URI).

```
1 <div xmlns:dc="http://purl.org/dc/elements/1.1/">
2    <div about="/alice/posts/trouble_with_bob">
3       <h2 property="dc:title">Learning Opportunity Provider</h2>
4       <p property="dc:description">Learning Opportunity Providers provide
             all kind of educational events. They include educational
            institutions but also instructors or experts.</p>
5    </div>
6 </div>
```

The underlying abstract representation of RDFa is RDF [1, Ch. 4]. This is shown by the `about` and `property` attributes in the example above. Therefore, it is possible to extract RDF triples from an RDFa annotated web page by a simple mapping [15, p. 2]. RDFa represents RDF structure with pure XHTML, can be built with any vocabularies, and is extensible. This allows new specifications for describing content of web pages. However, it does not ensure that content of the same area is described with the same specification.

Up to now, RDFa has been less implemented than microformats.

4.4.2 Microformats

Microformats have been developed by the people behind microformats.org[13] and are defined on their website as follows:

> Designed for humans first and machines second, microformats are a set of simple, open data formats built upon existing and widely adopted standards. Instead of throwing away what works today, microformats intend to solve simpler problems first by adapting to current behaviors and usage patterns (e.g., XHTML, blogging).

[13] http://microformats.org

This set of open data formats includes inter alia the rel and rev attributes that describe relationships of documents connected by links and their reverse links, XFN for describing relationships between people, hCalendar, hCard, and hAtom for content published as a web feed [4, Ch. 4–10].

Unlike RDFa, the technology of microformats can be integrated into both XHTML and HTML, and also works perfectly with CSS [15, p. 4]. It provides a compact syntax based on existing HTML, and is easy to implement. The following code shows an example using the hCalender microformat to describe an event.

```
1 <div id="hcalendar-event" class="vevent">
2   <abbr title="2009-07-01" class="dtstart">July 1th 2009</abbr>,
3   <abbr title="2009-07-03" class="dtend">July 3rd 2009</abbr>
4   <span class="summary">Learning Opportunity</span>
5   <div class="description">This Learning Opportunity is specified by its
        start and end date.</div>
6 </div>
```

A drawback of microformats is the limited number of specifications, as well as the difficult invention of new microformats. In addition, each microformat requires a separate parsing rule [15, p. 4].

4.4.3 Web Scraping for MLO-AD

Web scraping is a popular technique for retrieving different kinds of information on the Web. It can also be used to collect information about learning opportunities described by MLO-AD. However, to be able to extract data specified by the complex MLO-AD data model, it needs to provide semantics through inline metadata formats to identify the various properties of a learning opportunity and its provider.

Currently, microformats are more widely used and easier to implement than RDFa. The microformats technology is comprised of a limited set of existing specifications, whereas some of them can be used to describe properties of MLO-AD objects. The hCalendar microformat, which is a calendaring and events format, would be best applicable to describe educational events. It covers properties for the date, duration, location, category, and description, which are needed by a learning opportunity instance. The hCard microformat is a format to represent people, companies, and organizations. It provides properties that can be used to describe a provider of an educational event. However, the MLO-AD data model includes more properties than those provided by microformats, which are, for example, credit, qualification, prerequisite, and engagement. These properties cannot be described by existing specifications of microformats. Due to the fact that the specifica-

tion of a new microformat is a difficult process, microformats are not well applicable as an inline metadata format for learning opportunities specified by MLO-AD.

RDFa can only be used on XHTML web pages, but is easily extensible with new vocabularies. The extension through new concepts or properties results from the indication (by using simply a URL) of a directory, which is saved somewhere on the Web and from where the specific new properties are imported [1, Ch. 2]. This means that to include MLO-AD properties on a document annotated by RDFa, these vocabularies need to be specified and published on the Internet, so that references to these MLO-AD properties can be created by including this namespace. However, the provision of these vocabularies does not ensure that required information is included for the comparison of learning opportunities, or that the web page meets a specific structure for an easy and valid parsing of the data.

Web scraping of data described by MLO-AD is possible if the web page includes annotations with RDFa which identify the specific properties of learning opportunities and their providers. However, harvesters using the web scraping technique still have to face difficulties like missing or conflicting data or different structures of web pages describing learning opportunities.

4.5 Web Services

The World Wide Web Consortium defines web services as follows [16, Ch. 2]:

> A web service is a software system designed to support interoperable machine-to-machine interaction over a network.

Web services have an interface to access application functionality by using the HTTP protocol. A simple web service workflow consists of a request by a client (web service requester) to the server (web service provider), and the subsequent response from the server. The form, as well as the underlying protocols of the request and the response, depend on the type of the web service. The most common types of web services are *SOAP-based* and *RESTful*.

4.5.1 SOAP-based Web Services

SOAP-based web services (also called "big web services") use the *Simple Object Access Protocol* (SOAP) as an interoperable messaging format. A SOAP message is an XML document that follows the SOAP standard by

including three elements: an envelope, a header, and a body [40, Sec. 3.3.1]. SOAP is an envelope which contains any XML data, and is itself wrapped by the HTTP request or response.

SOAP-based web services usually use an architecture based on Remote Procedure Calls (RPC) [33, p. 19]. This means that to access a specific functionality of the server, the web service requester calls a function "remotely" on the server via the web service. SOAP-based web services have no limitation for defining functions, which is a main difference to RESTful web services.

The interface of the web service is described in a machine-processable format called Web Services Description Language (WSDL) [16, Ch. 2]. With WSDL, clients can find out how to use the specific web service and which functions can be called.

4.5.2 RESTful Web Services

REpresentational State Transfer (RESTful) web services have an architecture similar to that of the Internet, where each resource (web page in the web architecture) is uniquely addressable [33, p. 13]. Therefore, RESTful web services are resource-oriented, and not oriented to functions like RPC used by SOAP-based web services. Resources are identified by URIs.

In REST requests, the method information goes into the HTTP method and the resource information is the URI sent with the request [33, p. 13]. Due to the fact that HTTP methods are used, the following methods are the most important ones available for RESTful web services:

- GET retrieves or reads the resource with the given URI
- POST creates the resource with the given URI
- PUT updates the resource with the given URI
- DELETE deletes the resource with the given URI

RESTful web services support different representation formats. The most common formats are simple XML and JSON, but XHTML, RDF, or Atom can also be used [33, pp. 259–272].

Compared to SOAP-based web services, RESTful web services have gained more popularity because of their easier implementation.

4.5.3 Web Services for MLO-AD

Web services can be used for the communication between the data repository which holds information about learning opportunities and the aggregator portal based on MLO-AD. For this purpose, the learning opportunity provider must implement a web service which the aggregator portal can use to retrieve data about the provider's educational events. This implementation requires additional effort from the learning opportunity provider. However, the web service can also be used to provide other systems with information about learning opportunities.

Due to the fact that the communication between the aggregator portal and the provider is relatively simple, the web service does not need to consist of a complex implementation as SOAP-based web services would provide. RESTful web services are well applicable for retrieving information about learning opportunities and their providers. The aggregator portal uses the GET function with the URI of the specific learning opportunity object or of a collection holding more educational events and creates a new object or updates an existing one in its data repository. The RESTful web service of a learning opportunity provider can return various data types describing the educational event based on MLO-AD. Therefore, the aggregator portal must be able to parse responses of different types from the various learning opportunity providers. Regarding the structure of the MLO-AD data model, web services will usually return MLO-AD data in XML or RDF.

To keep the aggregator's data repository up-to-date, the aggregator must send the web service requests with the specific URIs at regular intervals. Unfortunately, the aggregator does not know if an update is actually necessary before sending the request. Even if the resource includes a timestamp of the last update, the aggregator portal can compare the dates just after receiving the response from the web service provider. This means that unnecessary data transfer can occur when using web services for collecting MLO-AD data.

4.6 OAI-PMH

The *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) is a protocol of the Open Archives Initiative[14] and provides an application-independent framework based on metadata harvesting [24, Ch. 1]. The protocol is based on HTTP and is used to make a digital repository's metadata available for harvest [32, p. 161].

[14] http://www.openarchives.org

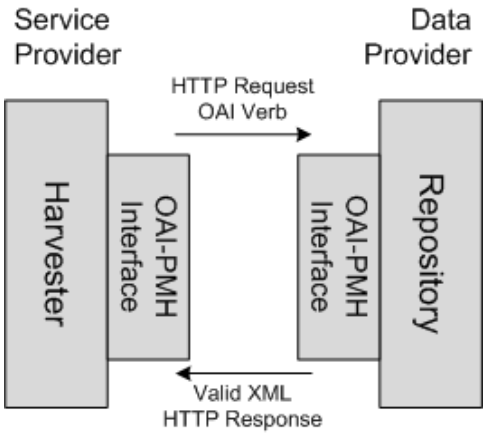


Figure 4.3: OAI-PMH Service Provider and Data Provider Architecture (according to [12, p. 71])

The architecture of OAI-PMH for harvesting information is build on two components: OAI data providers hold collections of content described by metadata, while OAI service providers harvest this content [12, p. 6]. Both components need an OAI-PMH interface to be able to communicate via the OAI-PMH protocol. Fig. 4.3 shows the architecture of OAI-PMH. Repositories of data providers are most commonly based on XML files, SQL databases, or Content Management Systems. With the Repository Explorer[15], data providers can test their archives for compliance with OAI-PMH.

All data providers must disseminate their metadata items in simple Dublin Core, but can additionally support other metadata formats [24, Sec. 3.4]. *Dublin Core* is a set of metadata elements that describe networked resources [12, p. 33]. It was first released 1998 by the Dublin Core Metadata Initiative (DCMI)[16].

4.6.1 OAI-PMH HTTP Requests and Responses

Requests for harvesting metadata via OAI-PMH are HTTP `GET` or `POST` requests. OAI-PMH defines six verbs to harvest metadata, whereas one, and only one, has to be included in the HTTP request. The finite list of verbs covers the following elements, as described in [32, pp. 162–168]:

- *GetRecord*: This verb returns an individual metadata record from the data repository. A metadata record covers all the information about

[15] http://re.cs.uct.ac.za
[16] http://dublincore.org

a given resource and is uniquely identified by its identifier, metadata prefix and datestamp. The GetRecord request requires the use of the identifier and the metadataPrefix arguments. The latter specifies the metadata schema of the OAI-PMH response.

- *Identify*: This verb is used to receive information about the repository, such as name, base URL, administrator's e-mail address, and version of OAI-PMH.
- *ListIdentifiers*: This verb retrieves the identifiers of a set of items from the data provider. This request must include a metadataPrefix and can contain arguments for a date range or a limit by set.
- *ListMetadataFormats*: This verb returns the metadata schemas supported by the OAI-PMH repository.
- *ListRecords*: This verb is used to retrieve a list of full metadata records. The metadataPrefix argument is required to receive a specific metadata preference. More arguments can be added to limit records by date or by set.
- *ListSets*: This verb returns information about the current list sets registered on an OAI-PMH server.

Responses of OAI-PMH requests are valid XML files conforming to the XML schema of OAI-PMH[17]. The workflow of OAI-PMH based on HTTP requests and responses is also displayed in Fig. 4.3.

4.6.2 OAI-PMH for MLO-AD

The main requirement of OAI-PMH in regard to its data providers is the dissemination of Dublin Core as the metadata format. Besides specific properties about learning opportunities, MLO-AD also includes the elements of Dublin Core in its "Learning Opportunity Object" class. Therefore, MLO-AD can fulfill this requirement, and is applicable for harvesting via OAI-PMH.

To be able to support MLO-AD, the data provider must include this metadata format via an XML schema, and by using a new namespace. This XML schema defines the specific elements and the structure of a metadata record about a learning opportunity resource.

An OAI-PMH harvester can only harvest metadata from a repository if this supports OAI-PMH via an interface. Therefore, providers of learning opportunities must implement a service to be able to accept OAI-PMH requests

[17] http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd

and return a valid XML file containing the required information about a learning opportunity or its provider. This implementation is an additional investment for an educational institution. However, with the provision of an OAI-PMH interface, the provider enables an easy and standardized way for harvesters to retrieve information about learning opportunities and to be able to advertise these educational events.

A harvester for MLO-AD resources can use the ListMetadataFormats verb in an HTTP request to ensure that the data repository supports the MLO-AD metadata format. After that, the harvester sends an HTTP request with the ListRecords verb to retrieve a list of full metadata records. The metadataPrefix argument must be added to the request, which specifies the MLO-AD as a preferenced metadata format. The optional arguments "from" and "until" limit the harvested list of metadata records by comparing the date stamp of the last modification of each metadata record with the argument values. This means that only updated or added learning opportunity resources from the last harvesting process are collected if the "from" argument with the specific date is added to the request.

4.7 Comparison of Solutions for MLO-AD

Every technology for collecting data has to be adjusted to the specification of the MLO-AD data model. For each technology, the learning opportunity provider has to provide a special implementation or mapping of MLO-AD data to enable data to be collected from its repository. This could be a creation and publication of Atom feeds, the mapping of MLO-AD information to the RDF specification, as well as enabling a remote query of these documents, an extension of the web pages containing the information about learning opportunities with RDFa annotations, the creation of RESTful web services or the implementation of an OAI-PMH interface. In the future, the aggregator portal based on MLO-AD will support all possible solutions for data collections to address a wide range of learning opportunity providers. However, the first version of the aggregator portal must support a way of data collection that is appealing to many educational institutions. This means that the investments to be able to support this way of data collection, as well as its administration, must be easy and cheap. Additionally, the chosen technology should also provide a secure and stable solution for the aggregator portal, so that it can parse the collected data easily and correctly. Proper data collection by the aggregator portal ensures a good provision of information about learning opportunities and an easy comparison of those.

RDF is a strong but complex specification to describe resources. However, it has difficulties establishing itself on the Web due to its complex realization.

Learning opportunity providers will prefer to describe their educational offers via XML documents including the MLO-AD specification, which are easier to create. The use of XML is very popular in the Web, so it can be assumed that the creation of such documents is well-known. Additionally, XML documents can include an XML schema which specifies the structure of the document. An XML schema specified for describing resources with MLO-AD can be published on the website of the aggregator portal, so that the educational institutions can include it in their XML documents. Of course, it cannot be ensured that learning opportunity providers actually use the schema. However, if the XML schema is included, the parsing of the XML document is facilitated for the aggregator portal.

Alternatively, learning opportunities could also be described directly on a web page, but must include RDFa annotations to enable web scraping for the aggregator portal. It is easier to include RDFa in a website than to create an RDF document to describe resources based on MLO-AD. Although RDFa is less well-known than XML, it is gaining more and more popularity due to its easy implementation. However, it is not possible to define a structure of the provided content on the web page, which complicates a proper mapping of data to the properties of MLO-AD. Additionally, like XML and RDF, RDFa is not a technology to transmit information over the Web, but it is a good and easy solution to display and save data about learning opportunities based on MLO-AD.

Web services are a popular technology to communicate between web applications and to transfer information described by XML. RESTful web services are especially easy to create and use. However, the web service solution has the drawback of unnecessary data transfer, as it does not allow the last modification date of a resource to be checked before it is received.

By using Atom feeds for collecting data about learning opportunities, the aggregator portal can check whether this data transfer is actually necessary before the content is downloaded. Feeds are widely used over the Web and easy to create, especially as many tools which facilitate the creation and publication of feeds already exist. Unfortunately, feeds have the same disadvantage as annotations of RDFa: Web feeds do not include a definition of the content structure which describes the learning opportunity. Because of this, the parsing of the XML content can be error-prone. However, due to Atom's recognition and popularity, this technology is still a good solution for collecting data from learning opportunity providers.

The technology via OAI-PMH returns XML files that can include an XML schema for describing the structure of the document. It is also possible to avoid unnecessary data transfer due to the valid request parameters which allow a restriction of the retrieved content according to the date of last

modification. However, OAI-PMH also requires a special implementation of an OAI-PMH interface to accept HTTP requests based on this protocol. A downside of this technology in connection with the implementation is the fact that OAI-PMH is not well-known and institutions may be reluctant to use it. However, a solution based on OAI-PMH is easy for the learning opportunity provider to manage and is also secure when harvesting and parsing for the aggregator portal.

Recommended Solution

The recommended solution for MLO-AD is developed by using OAI-PMH, because of its stability. To enable data collection via OAI-PMH, the aggregator portal and all learning opportunity providers have to implement an interface which can understand this protocol. The OAI-PMH interface of the educational institution depends on the system of their data repository, as the interface has to create the XML document including the information from the repository. If the system of the data repository is known and commonly used, an OAI-PMH interface can be provided to enable easy and automatic data provision for the learning opportunity provider. After the installation of this component, the educational institution can easily accept HTTP requests based on OAI-PMH and send a valid XML file as a response that includes the information about learning opportunities.

The XML file containing MLO-AD information is enriched by a namespace of this standard that guarantees the use of valid elements. Additionally, it includes an XML schema ensuring that the structure of the content matches the data model of MLO-AD. The schema also specifies the occurrences of each element and can define default values. Therefore, the XML schema guarantees that harvested MLO-AD content complies with the requirements of the aggregator portal.

Course management systems like Moodle[18] and Sakai[19] are widely used by educational institutions, and usually cover all the information about learning opportunities provided by the institution. An OAI-PMH interface which can communicate with a Course Management System is definitely appealing to learning opportunity providers who want to provide their information for the aggregator portal based on MLO-AD. Therefore, the first solution of MLO-AD data collection will additionally provide an OAI-PMH interface that can be integrated with the Course Management System Sakai. Learning opportunity providers can download this component from the website of the aggregator portal and install it in their Sakai system.

[18] http://moodle.org
[19] http://sakaiproject.org

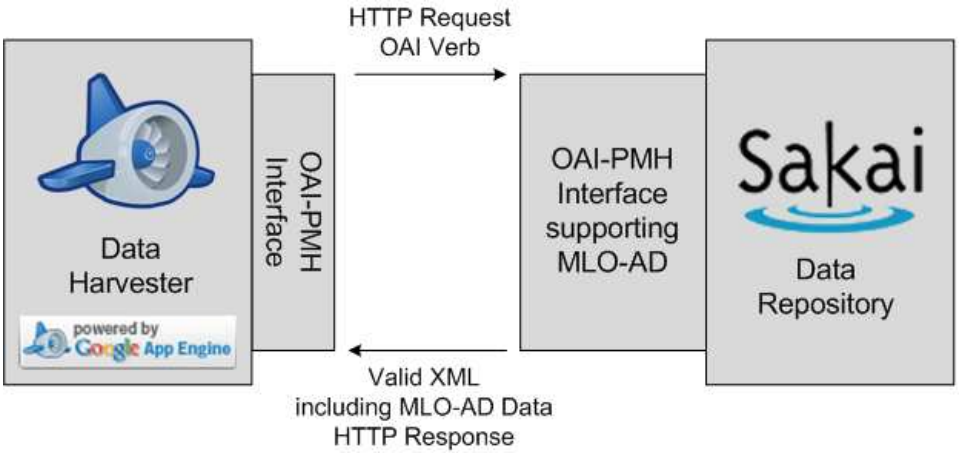


Figure 4.4: System architecture of MLO-AD harvester

Fig. 4.4 shows the architecture of the MLO-AD data collection based on OAI-PMH. The aggregator portal is developed by using the Google App Engine[20], which enables an easy build of web applications on a scalable system. The Google App Engine supports Java as its development language.

[20] http://code.google.com/appengine

Chapter 5

User Interface

The aggregator portal based on MLO-AD provides two different kinds of user interfaces: The *web portal* is a web page that is accessible via a public URL, while the *social networking gadget* provides integration with a social network service. Both types use the same logical system as back-ends and have access to the same data about learning opportunities. The interfaces differ in the possibilities of interaction, as well as the arrangement of web elements and information. They are adjusted to the user experience and the applicability of web pages and social network services.

5.1 Web Portal

The web page of the aggregator portal allows prospective learners to browse and search for learning opportunities. Additionally, it provides general information and services about the system. It is the main reference to the aggregator portal based on MLO-AD.

Besides the data about learning opportunities, the web portal offers information for learning opportunity providers to enable them to advertise their educational events via the aggregator portal. Educational institutions and instructors can register and sign in to get to know the technologies that the portal supports for collecting data based on MLO-AD. Explanations and tutorials help learning opportunity providers to integrate MLO-AD into their system and to implement their chosen technology to enable data collection. To facilitate the implementation, the portal also provides certain data repository tools and interfaces, which can be downloaded from the web page. Educational institutions can also find a form to submit or manage their information manually.

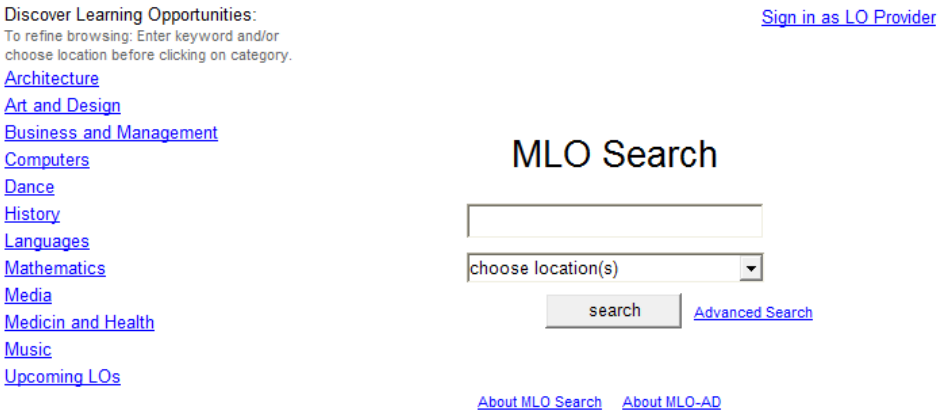


Figure 5.1: Wireframe of the web portal's main page

Although the web portal offers much information and interaction, the interface has to remain simple to provide easy orientation. Therefore, the main page of the portal displays search fields in the center of the page, and categories for browsing through learning opportunities on the left side. The arrangement of elements for the main page of the web portal is shown as a wireframe in Fig. 5.1. Wireframes are commonly used in interface design to provide a visual guide to the structure of an interface and the relationships between its pages before the design is developed [10, p. 91].

5.1.1 Searching for Learning Opportunities

The search area of the web portal includes an input field for the subject or title of the learning opportunity and a drop-down menu for choosing a location. The selection of the location is optional, but can also consist of multiple values because of its multi-select functionality.

To be able to add more specified information about a learning opportunity to a search query, users can switch to another page by clicking on "Advanced Search." This view includes input fields for properties like start date (supported by a calendar to pick a date), qualification, cost, duration, credits, language of instruction, engagement, prerequisite, assessment, and again location. After inserting values for one or more of these properties, the portal creates a combined query which is set into the main input field of the search mask. Advanced users can directly input such a combined query without changing the view to the advanced search mode. The following line shows an example of a combined query for the MLO-AD web portal.

```
1 java programming location:montreal cost:<500 prerequisite:none
```

This feature allows a functional refinement of a search for both common and advanced users, and keeps the main page of the search screen simple and clear.

5.1.2 Display of Search Results

By default, the results of a search request are grouped by the educational institution or instructor, who is the learning opportunity provider. To save the prospective learner from a flood of information, the survey of the results only shows the name and location of each provider with the name of a learning opportunity, as well as the starting dates of its sessions. The names of the institution and its offers act as a link to the source of this information (usually the web page of the institution). If the search request returns more than three educational events per provider, the number of founded offers is shown instead of their names. With a click on this number, all learning opportunities of the institution are displayed. The general view of results works with disclosure panels, which can be opened and closed to show and hide further information about a learning opportunity or its provider. Fig. 5.2 illustrates a wireframe with search results in the general display mode.

To respond to different needs of users and to facilitate their search, the results can be sorted and grouped by different properties. The button "Sort By" opens a menu that includes program type, qualification, cost, starting date, location, duration, credit, and engagement. The results are displayed in the same way as in the general display mode, except that the "header" of a result (which is the learning opportunity provider in the default view) is replaced with the specific property.

Certain properties can be better compared by a design that is different from the display of simple text. Therefore, the web portal allows its view of search results to be switched to a *map* or a *calendar*. A map similar to a Google Map gives a good survey of the location of learning opportunities. Pins mark an educational event on the map, and show further information when the mouse pointer is moved over them. On the contrary, the display of search results on a calendar facilitates the search and comparison of events for prospective learners with a specific time frame. Colored fields indicate sessions of learning opportunities. By integrating the duration and engagement (specifies the pattern of attendance such as evenings, or daytime), the calendar can show the complete time of each educational event that the learner has to expend on it. Different colors of fields identify different learning opportunities. Some of them can also be held on multiple dates. The calendar display mode can show multiple sessions of an educational event by fields with the same color, but different intensity.

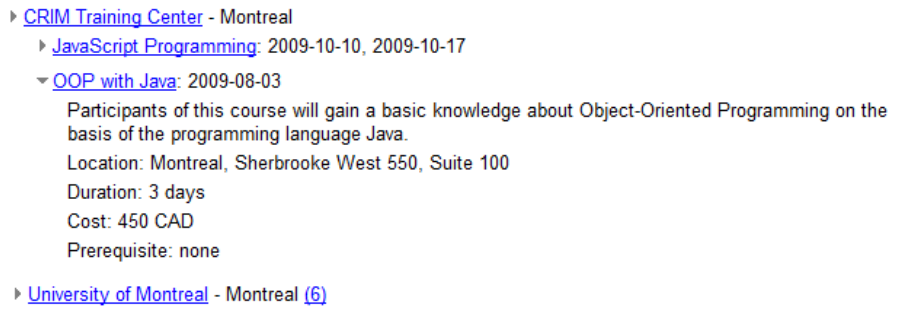


Figure 5.2: Wireframe showing the general display mode of search results

5.1.3 Browsing through Learning Opportunities

Users can discover learning opportunities by browsing through those available on the aggregator portal based on MLO-AD. The main page of the web portal displays categories of educational events (see Fig. 5.1). Many comprehend sub-categories, which are shown underneath their parent after they are clicked upon. Learning opportunities are displayed in the same way as the results of a search request. It is also possible to change the view according to the location or the date, and to sort the events by different properties.

In addition, the web portal allows the browsing results to be refined exactly as when refining a search. Before clicking on a link of a category, a prospective learner can choose one or more locations or enter keywords into the main input field. Hence, the user can specify requirements, but still discover educational events that could not be found by a search request.

The area that comprises the categories, as well as each sub-category, includes an additional link called "Upcoming LOs." This feature allows irresolute learners to discover learning opportunities (that belong to one specific or to all kind of categories) starting within the following days.

5.1.4 Comparison of Learning Opportunities

Besides the different modes to display search and browsing results, the web portal provides another feature for a better comparison of learning opportunities. Users can find a checkbox next to each result to add an educational event to a comparison matrix. After the selection and the click on the button "Create LO Matrix", another page for choosing MLO-AD properties that should be included in the matrix is displayed. The created matrix shows the chosen learning opportunities as rows, and the properties of MLO-AD as columns, and provides a good survey of the events the user is interested in.

Figure 5.3: Wireframe of the social networking gadget

5.2 Social Networking Gadget

Social network services enable new functionalities for gadgets due to the different user behavior in these systems. Social networking means communicating with friends or people who share the same interests. Popular gadgets enable this communication, and are integrated in common actions within social network services.

The social networking gadget of the aggregator portal based on MLO-AD is a miniature of the web portal. Therefore, it offers only a limited functionality of that, but it is adjusted to the user experience of social network environments.

The gadget provides the search functionality for learning opportunities in nearly the same way as the web portal. In a social networking environment, only the main input field is displayed to allow a simple or combined search query to be entered. The advanced search is still enabled, but without providing a different view to input properties in specific fields. Instead, the gadget offers a "Help" button which opens a small note that explains the creation of a combined search query. Results of a search are shown like results in the web portal (by default, in the general display mode). They can also be sorted and grouped by different properties. The gadget does not provide a display

of search results on a map or calendar. Browsing learning opportunities or comparing results within a matrix is not supported. However, a link to the web portal is always visible, so that users can easily access all the functionality of the portal. Fig. 5.3 shows the structure of the social networking gadget with a wireframe.

From the technical point of view, gadgets are XML files that include HTML for markup, JavaScript for interactivity, and CSS for presentation. They can also include an inline-frame to point to another web page and represent its content. The gadget for the aggregator portal based on MLO-AD uses the OpenSocial[1] API for communicating with the social network service. Social applications based on OpenSocial can be integrated with all social network services that support this API. Functions of OpenSocial allow, among others, information about the current user and friends to be received, activities that describe what the user is doing to be modified, and messages to contacts of the user to be sent.

The social networking gadget for MLO-AD allows users to share learning opportunities with their friends. The "Share with Contacts" button next to each search result opens a window with a list of the user's friends. Accordingly, the user can choose one or more contacts by clicking their names in the list. Hence, the system notifies these contacts of the educational event by sending a message (within the social network service — this feature is provided by all services that support OpenSocial) including the information about the event.

Many users of social network services like to inform their friends about their current activities. A session of a learning opportunity found by the MLO-AD gadget can also be referenced as an activity. A click on the button opens a window where users can indicate whether they will or have already attended this event or whether they are interested in an event but cannot participate because of a certain reason. This feature also enables the adding of a comment to the activity which includes the learning opportunity. After this process, a link to the learning opportunity, the user's opinion of it, as well as the comment, are shown as the current activity of the user.

Most social network services already allow the creation of a group that everyone or only certain people can participate in. A person interested in a learning opportunity or the instructor of the event can create a group relating to with it by clicking on the button. Accordingly, friends can be invited to this group, discussions can be started, questions can be asked, and comments can be posted.

[1] http://www.opensocial.org

Chapter 6

The Prototype

The first web application based on MLO-AD places its emphasis on the implementation of data collection. It implements the recommended solution in Section 4.7 for aggregating MLO-AD data. Additionally, a simple user interface is provided to allow searches for learning opportunities.

The following three components form the architecture of the prototype: The *aggregator portal* is the core element of the implementation. It collects data about learning opportunities and provides the aggregated information for users or other services. The second part is the *Sakai plug-in*, which facilitates the support of data collection for learning opportunity providers, while the third element, the *OpenSocial gadget*, allows an integration of the portal's features with social networking services.

The implementation's aim is to meet the success criteria of the prototype that are described in Section 3.3. In a nutshell, the prototype aims to demonstrate the possibilities and applicabilities of the MLO-AD standard.

6.1 Aggregator Portal with Google App Engine

The aggregator portal based on MLO-AD is the connector between multiple providers of learning opportunities and prospective learners or services offering information about learning opportunities. Its main target is the collection of data structured by MLO-AD. With a database full of aggregated learning opportunities, the portal can supply learners with their needed information in a structured and comprehensive way.

The portal is developed with the programming language Java on Google App Engine (GAE)[1]. This complete development stack offers familiar technologies

[1] http://code.google.com/appengine

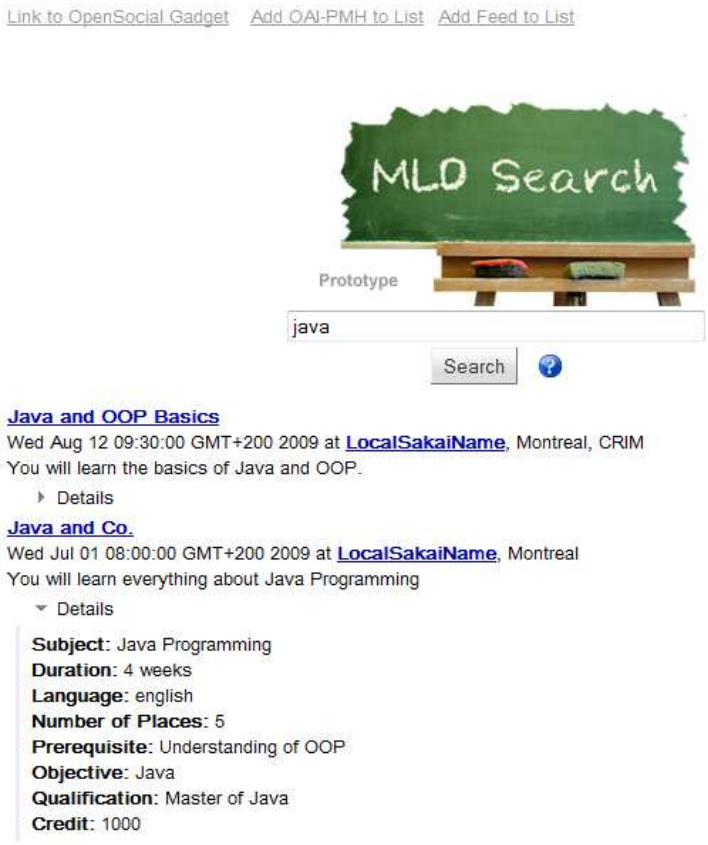


Figure 6.1: Aggregator portal of the prototype

to build and host web applications that can be easily maintained and scaled. The maintenance of the server is carried out by Google.

The front-end of the prototype is created with Google Web Toolkit (GWT)[2]. This technology allows a highly performance JavaScript front-end including AJAX to be easily built. The code is written in the programming language Java, which is converted to JavaScript by GWT, and works across all major browsers. Fig. 6.1 shows the portal's front-end created with GWT.

For the database, the prototype uses Java Data Object (JDO)[3] that makes the data of learning opportunities and their providers persistent. JDO works with annotations on Java classes ("plain old Java objects" or POJOs). These annotations define how to save instances of the class and which of its properties. Instances are stored as entities in the database of App Engine. JDO includes a query interface (JDOQL) that allows an easy recreation of class

[2] http://code.google.com/webtoolkit
[3] http://java.sun.com/jdo

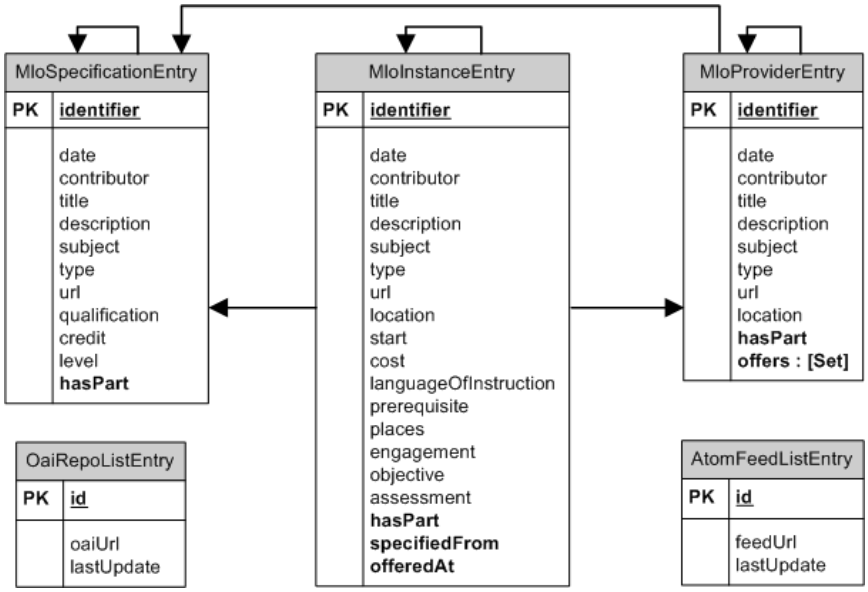


Figure 6.2: The prototype's database model

instances from the entities of the database. The prototype's database consists of various entity types. For the properties of MLO-AD, each resource of the standard's data model (Learning Opportunity Provider, Learning Opportunity Specification, and Learning Opportunity Instance) forms a single entity type. Instances of these types are connected with properties like `offers`, `offeredAd`, and `specifiedFrom`. Additionally, each MLO-AD object can refer to another instance of its type by using the `hasPart` reference. The database also includes other entities needed to save information for collecting data from learning opportunity providers. The complete database model of the prototype is shown in Fig. 6.2.

The collection of MLO-AD data is implemented with the technology recommended in Section 4.7. Therefore, the prototype includes an OAI-PMH harvester for MLO-AD metadata. Additionally, the portal provides an aggregator for Atom feeds that contain MLO-AD information to show a second suitable possibility for collecting information based on this specific standard. The following subsections describe the implementations of both solutions.

The prototype of the aggregator portal based on MLO-AD is published at http://mlo-app.appspot.com.

6.1.1 OAI-PMH Harvester

The portal's harvester collects information about learning opportunities via OAI-PMH from repositories that support this harvesting protocol. Links to these repositories are saved in the `OaiRepoListEntry` table of the database (see Fig. 6.2). OAI-PMH works with simple HTTP requests and six different verb parameters to define the action of the request. The OAI-PMH harvester of the MLO-AD portal uses two verbs for collecting and updating its data: *ListMetadataFormats* returns a list of the repository's supported metadata formats and confirms the support of MLO-AD. The actual data about learning opportunities are delivered via the *ListRecords* verb, and can be limited to MLO-AD data by adding a metadata prefix as a parameter.

The response of the OAI-PMH request is an XML file, which is read by using the de facto standard Simple API for XML (SAX). SAX is an event-based parser that reads the document in real time [18, Ch. 20]. Therefore, the processing can start before the entire document is read, which saves time and, most importantly, memory.

After the first OAI-PMH request, the MLO-AD harvester reads the returned XML file (which contains the supported metadata formats) and checks if MLO-AD is included. If so, the request with the `ListRecords` verb is sent with attached parameters for the metadata prefix of MLO-AD, the date of the last update from `OaiRepoListEntry` as the `from` parameter, and the current date as the `until` parameter. The date parameters ensure that only new or updated OAI-PMH records are returned. Like the first response, the newly received XML file is parsed by the SAX parser and the information about learning opportunities is saved in the App Engine database. Due to the fact that this XML file is referenced to an XML schema (whose reference is also included in the `ListMetadataFormats` response), parsing MLO-AD records from an OAI-PMH repository is relatively error-free.

6.1.2 Atom Feed Aggregator

The second technology for collecting MLO-AD data is based on Atom feeds. This portal includes an aggregator that parses Atom feeds and saves the information described by MLO-AD properties in the database.

Links to the feeds, which are aggregated at regular intervals, are saved as a feed list in the database. This list is called `AtomFeedListEntry` (see Fig. 6.2), and includes the date of the last update for each entry. For an update of the data, the aggregator iterates this list and compares the date in the list with the update date of the Atom feed, which must be included. Therefore, only feeds that were modified since the last update are parsed.

To parse an Atom feed, the aggregator portal uses ROME[4], which is a set of open source Java tools for parsing, generating and publishing RSS and Atom feeds. It is based on the JDOM[5] XML parser to generate Java objects from the feed's XML source. ROME was chosen for the aggregator portal not only because of its enhancement of the parsing process, but also because of its easy integration into Google App Engine. Although ROME facilitates reading an Atom feed, the content of the feed that includes XML tags with MLO-AD properties must be parsed "manually." The aggregator receives a String element of the feed's content, searches for each valid property in this, and saves the appendant value in the database.

6.1.3 Automatic Update of MLO-AD Data

Google App Engine includes a Cron Service to execute scheduled tasks at specific times or regular intervals. These so-called *cron jobs* are commonly used in Unix-like computer operating systems for an automatic update of data or the creation of backup files[6]. In App Engine, the `cron.xml` file defines cron jobs with specific URLs that are invoked at a given time of day.

The aggregator portal based on MLO-AD needs scheduled tasks to collect regularly information about learning opportunities and to keep the data up-to-date. The following lines show the `cron.xml` of the prototype. It contains a cron job that updates the data of all Atom feeds every day at midnight New York time zone. Every day at 12:30 a.m. New York time zone, the second cron job invokes an URL to harvest and update data from all known OAI-PMH repositories.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <cronentries>
3   <cron>
4     <url>/mloapp/mloAtom</url>
5     <description>Updates data of all Atom feeds from the feed list.</
          description>
6     <schedule>every day 00:00</schedule>
7     <timezone>America/New_York</timezone>
8   </cron>
9   <cron>
10    <url>/mloapp/mloOai</url>
11    <description>Updates data of all referenced OAI-PMH repositories.</
          description>
12    <schedule>every day 00:30</schedule>
13    <timezone>America/New_York</timezone>
14  </cron>
15 </cronentries>
```

[4] https://rome.dev.java.net
[5] http://www.jdom.org
[6] http://www.unixgeeks.org/security/newbie/unix/cron-1.html

6.1.4 Search Functionality

The prototype includes only one input field to search for learning opportunities. Although the user interface of the prototype does not cover the possibility of switching to an advanced search view, the search allows a small version of an advanced search due to the integration of Compass[7]. Compass is a Java search engine framework developed as an open source project and built on the top of Lucene[8], which is an Apache project for search software. Like JDO, Compass uses annotations to define the searchable properties of a plain old Java object (POJO). Compass indexes the data to provide search results in a fast way. Another reason for choosing Compass as the prototype's search engine framework is its easy integration with Google App Engine.

For now, the prototype only allows searches for properties of a learning opportunity specification. Only this Java class is enriched with Compass annotations. In particular, the properties `title`, `subject`, and `type` are searchable with the prototype. The syntax of a combined search query looks as follows, whereas the last three parts are optional.

```
1 [keyword] title:[keyword] subject:[keyword] type:[keyword]
```

Search results, which are displayed underneath the search field as in Fig. 6.1, are shown in a simple way. A sorting of results by specific properties is not supported by the prototype. In addition, the possibility to change the view from text display to a map or calendar has not been implemented yet.

6.2 Sakai Plug-In

Many educational institutions use a course management system like Sakai[9] or Moodle[10] for managing their courses and resources. The prototype provides a Sakai plug-in (developed for version 2.5.4) to facilitate the provision of information based on MLO-AD for institutions using Sakai.

Sakai is built with Java and the Spring framework[11], and consists of different tools for coursework, communication, and collaborative work. New components can easily be added as tools or as simple Java web servlets.

To support both data collection technologies of the prototype's aggregator, the Sakai plug-in includes an interface for OAI-PMH and a feed generator

[7] http://www.compass-project.org
[8] http://lucene.apache.org
[9] http://sakaiproject.org
[10] http://moodle.org
[11] http://www.springsource.org

for Atom feeds with MLO-AD properties. Both features are implemented as Java web servlets and can be invoked via their servlet URL pattern. Sakai has different possibilities to describe properties of courses or the institution, which are, for example, course site descriptions, as well as general or config information in a properties file. However, not all MLO-AD-required information can be supported by Sakai. Therefore, the Sakai plug-in additionally contains a template for describing courses, which is enriched with RDFa annotations. This template is easy to manage and also demonstrates the practicability of RDFa.

6.2.1 RDFa Template

Each course in Sakai uses a specific Sakai site, which can contain different tools such as syllabus, resources, calendar, portfolio, and chat. Sites in Sakai can also include a general site description to give information about the course. Sakai provides different templates for these descriptions, which can be used for a specific structure of the information.

The Sakai plug-in of the aggregator portal attaches a new site description template to the available ones of Sakai. This template is a simple XHTML table with all valid properties of MLO-AD, whereby the user can easily add the specific values to the fields. The code of the table is enriched with RDFa annotations that provide semantics for machines. Therefore, the template does not only facilitate the support of MLO-AD information for learning opportunity providers, but also allows the extraction of MLO-AD data from this site for web bots or aggregators. The following lines illustrate an example of the template's source code.

```
1 <table width="100%" cellspacing="0" cellpadding="0" border="0" xmlns:xsd
      ="http://www.w3.org/2001/XMLSchema" xmlns:dc="http://purl.org/dc/
      elements/1.1/" xmlns:mlo="http://mlo-app.appspot.com/mlo/elements/"
      about="">
2   <tbody>
3     <tr><td>Title</td><td property="dc:title">Java and OOP Basics</td></
          tr>
4     <tr><td>Subject</td><td property="dc:subject">Java</td></tr>
5     <tr><td>Description</td><td property="dc:description">You will learn
           the basics of Java and OOP.</td></tr>
6     <tr><td>Contributor</td><td property="dc:contributor">Katharina</td
          ></tr>
7     <tr><td>Qualification</td><td property="mlo:qualification"> </
          td></tr>
8     <tr><td>Level</td><td property="mlo:level"> </td></tr>
9     <tr><td>Credits</td><td property="mlo:credit"> </td></tr>
10    <tr><td>Location</td><td property="mlo:location">Montreal, CRIM</td
          ></tr>
11    <tr><td>Start Date</td><td datatype="xsd:dateTime" property="mlo:
          start">2009-08-12 09:30:00</td></tr>
```

```
12     <tr><td>Duration</td><td property="mlo:duration">6 h</td></tr>
13     <tr><td>Cost</td><td property="mlo:cost">250 CAD</td></tr>
14     <tr><td>Prerequisite</td><td property="mlo:prerequisite">some</td></
           tr>
15     <tr><td>Max. Number of Places</td><td property="mlo:places">15</td
           ></tr>
16     <tr><td>Language of Instruction</td><td property="mlo:
           languageOfInstruction">english</td></tr>
17     <tr><td>Engagement</td><td property="mlo:engagement">daytime,
           workplace-based</td></tr>
18     <tr><td>Objective</td><td property="mlo:objective">Java</td></tr>
19     <tr><td>Assessment</td><td property="mlo:assessment"> </td></tr
           >
20   </tbody>
21 </table>
```

6.2.2 OAI-PMH Interface

To be able to accept OAI-PMH requests, Sakai needs an interface that prepares its data to support this protocol. Since implementations for OAI-PMH repositories already exist, the Sakai plug-in uses the open source framework OAICat[12] developed by the Online Computer Library Center (OCLC)[13] under the OCLC Research Public License 2.0. OAICat significantly facilitates the support of OAI-PMH, but still needs slight adjustments to MLO-AD and the data system of the learning opportunity provider or Sakai.

MLO-AD can be added to the OAICat framework by specifying its prefix, namespace and a valid XML schema. Due to integrated Dublin Core elements of MLO-AD, this standard is a valid metadata format for OAI-PMH. The Sakai plug-in takes the MLO-AD XML schema[14] created by Mark Stubbs from the MLO-AD working group as a reference.

OAICat already supports different data systems such as file system, database, and Java Database Connectivity (JDBC). However, it can always be extended by individual implementations to support another data repository. The prototype's OAI-PMH interface stays simple and uses the file system as the source of its data. It takes the information about learning opportunities and their providers from XML files and converts them for the transfer via OAI-PMH. These XML files comply with the specifications of MLO-AD's XML schema. The OAI-PMH interface of the prototype does not use the information of Sakai yet, but will meet this need in the future.

[12] http://www.oclc.org/research/software/oai/cat.htm
[13] http://www.oclc.org
[14] http://wiki.teria.no/confluence/display/CIF/MLO-AD+illustrative+XML+binding

6.2.3 Atom Feed Generator

For the second technology of data collection, the Sakai plug-in uses an Atom feed generator that creates or updates a feed whenever a modification of a site description including the MLO-AD template with RDFa annotations occurs. This happens via a JavaScript function which is attached to Sakai's site description and invokes the URL of the Atom generator servlet. Therefore, MLO-AD information in Sakai is always directly written in an Atom feed.

The generator reads the information of the site description and uses Apache Abdera[15] to generate the feed. Abdera is an open source Atom implementation for the IETF Atom Syndication Format and Atom Publishing Protocol standards. Properties of MLO-AD are written within XML tags in the `content` element of a feed entry. To define these properties as MLO-AD elements, the tags include an XML namespace extension of MLO-AD. Information of the learning opportunity provider is specified by children of the `feed` element. Each `entry` of a feed describes a learning opportunity specification including information about one or more appendant learning opportunity instances. The following feed is an example that contains one entry specifying a learning opportunity.

```
1 <?xml version='1.0' encoding='UTF-8'?>
2 <feed xmlns="http://www.w3.org/2005/Atom">
3   <id>tag:gtn.sakaiquebec.org,2009:/mlo-rdfa/mlo.xml</id>
4   <title type="text">GTN Quebec Sakai</title>
5   <author><name>GTN Quebec</name></author>
6   <contributor><name>GTN Quebec</name></contributor>
7   <link href="http://gtn.sakaiquebec.org/portal" rel="alternate" />
8   <link href="http://gtn.sakaiquebec.org/mlo-rdfa/mlo.xml" rel="self"/>
9   <updated>2009-06-22T19:00:01.293Z</updated>
10  <entry>
11    <id>tag:gtn.sakaiquebec.org,2009:/portal/site/1d9c9fbc-4ded-4755-82
          fd-1a58b70c04b6</id>
12    <title type="text">Java and OOP Basics</title>
13    <updated>2009-06-20T16:10:55.948Z</updated>
14    <published>2009-05-18T04:30:23.579Z</published>
15    <link href="http://gtn.sakaiquebec.org/portal/site/1d9c9fbc-4ded
          -4755-82fd-1a58b70c04b6"/>
16    <contributor><name>Katharina</name></contributor>
17    <summary type="text">You will learn the basics of Java and OOP.</
          summary>
18    <category term="Java" />
19    <content type="application/xml">
20      <dc:subject xmlns:dc="http://purl.org/dc/elements/1.1/">Java</dc:
            subject>
21      <mlo:location xmlns:mlo="http://mlo-app.appspot.com/mlo/elements/">
            Montreal, CRIM</mlo:location>
```

[15] http://abdera.apache.org

```
22      <mlo:start xmlns:mlo="http://mlo-app.appspot.com/mlo/elements
            /">2009-08-12T07:30:00.000Z</mlo:start>
23      <mlo:duration xmlns:mlo="http://mlo-app.appspot.com/mlo/elements
            /">6 h</mlo:duration>
24      <mlo:cost xmlns:mlo="http://mlo-app.appspot.com/mlo/elements/">250
            CAD</mlo:cost>
25      <mlo:languageOfInstruction xmlns:mlo="http://mlo-app.appspot.com/
            mlo/elements/">english</mlo:languageOfInstruction>
26      <mlo:places xmlns:mlo="http://mlo-app.appspot.com/mlo/elements
            /">15</mlo:places>
27      <mlo:engagement xmlns:mlo="http://mlo-app.appspot.com/mlo/elements
            /">daytime, workplace-based</mlo:engagement>
28    </content>
29  </entry>
30 </feed>
```

6.3 OpenSocial Gadget

The prototype implementation of the aggregator portal based on MLO-AD additionally covers an XML file which is a social gadget for social networking services. It can be integrated with all services that support the OpenSocial API, which are, among others, MySpace[16], LinkedIn[17], Xing[18], and Hi5[19].

This gadget is developed according to the XML schema of OpenSocial. The code below shows the full XML document of the current version. For the prototype, the gadget is a simple implementation that allows learning opportunities to be searched, but does not include social features like sharing educational events with friends. Therefore, the OpenSocial XML file only contains an inline-frame, pointing to a web page of the aggregator portal that is adjusted to the view of the social gadgets. An OpenSocial gadget can also specify CSS for the presentation, which would be defined by the `style` tag at line 14. The `script` tag from lines 15 to 17 contains the JavaScript code, which usually covers OpenSocial JavaScript functions to allow communication with the social environment. The implementation of the prototype only adjusts the height of the gadget and does not support social activity yet.

```
1 <?xml version="1.0" encoding="UTF-8" ?>
2 <Module>
3   <ModulePrefs title="MLO Search"
4       description="MLO Search allows searching for learning opportunities
            based on MLO-AD"
```

[16] http://www.myspace.com
[17] http://www.linkedin.com
[18] http://www.xing.com
[19] http://www.hi5.com

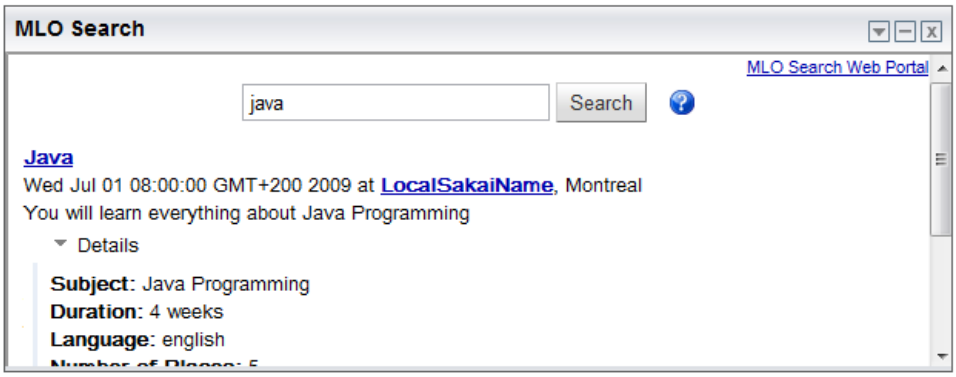


Figure 6.3: Screenshot of the prototype version of the OpenSocial gadget

```
5       author="Katharina Bauer-Oeppinger">
6     <Require feature="opensocial-0.8"/>
7     <Require feature="views" />
8     <Require feature="dynamic-height" />
9   </ModulePrefs>
10  <Content type="html" view="default,home,profile,canvas"> <![CDATA[
11  <iframe src ="http://mlo-app.appspot.com/MLOAppGadget.html" width
        ="100%" height="100%">
12    <p>Your browser does not support iframes.</p>
13  </iframe>
14  <style type="text/css"></style>
15  <script type="text/javascript">
16    gadgets.window.adjustHeight();
17  </script>
18  ]]>
19  </Content>
20 </Module>
```

After the search process, the results are displayed within the gadget in a simple way. The prototype does not support a sort of the results by specific properties. Fig. 6.3 shows the prototype gadget with search results.

Chapter 7

Conclusions

This thesis results in the analysis of data collection techniques specialized for the MLO-AD standard, the exploration of the fundamental design of a portal that offers learning opportunities, as well as the implementation of a prototype.

Different technologies, particularly those already used by existing aggregator systems, are analyzed according to their usability for collection of MLO-AD data. Due to the complex structure of MLO-AD, information based on this standard has to be enriched with semantics by using XML or a Semantic Web technology. The analysis shows that XML is used in many solutions (web feeds, web services, and OAI-PMH), due to the fact that it is the prevalent standard for exchanging data between applications over the Web. Additionally, it is well-known, commonly used, and can specify its structure with an XML schema. To comply with MLO-AD, the structure of the XML must match the data model of this standard. This can be ensured by a schema like the one created by Mark Stubbs from the MLO-AD working group[1]. The analyses of Semantic Web technologies (RDF, OWL, RDFa, and microformats) consider the possibilities for enhancing MLO-AD information with semantics to make it machine understandable. Although these technologies are strong languages for describing MLO-AD content, they are limited in the actual transfer of the data from the learning opportunity provider to the aggregator portal. Technologies for transferring RDF exist (RSS 1.0 and SPARQL), but they have not been further described in this thesis due to the fact that they are not commonly used.

Although this thesis analyzes the possibilities to describe and transfer MLO-AD data in detail, it does not consider the content of the data itself. The MLO-AD specification does not address vocabularies needed to ensure an

[1] http://wiki.teria.no/confluence/display/CIF/MLO-AD+illustrative+XML+binding

easy comparison of learning opportunities and an interoperability between different educational domains. Especially for properties like type, qualification, level, and engagement, predetermined values would make learning opportunities and their providers easily comparable. Additionally, the specification of MLO-AD does not always clearly declare the meaning of properties. For example, the date property (which is included in learning opportunity providers, specifications, as well as instances and is a different date than the start date of an instance — see Fig. 2.1 of Section 2.2) can be interpreted as the date of the last update from the provider, of the last update from the collector or of the resource's creation. The reason for not addressing vocabularies is that these values must be updated frequently. Therefore, separate CEN Workshop Agreements (CWAs) will deal with vocabularies for MLO-AD properties [39, pp. 2–3]. Guidelines for values of MLO-AD properties would facilitate the use of the standard and the comparison of learning opportunities and their providers.

Considering the design exploration of the aggregator portal based on MLO-AD, this thesis offers different and contemporary ideas for presenting the content or interacting with the system. Thereby, it specially addresses the needs and expectations of the target audience, which is analyzed in Section 3.1. Additionally, the exploration responds to the internet consumer phenomenon of social networking with an integration of the aggregator portal with such a system. One issue that is not considered in the design analysis is the fact that a learning opportunity can have a reference to another one, which is identified by the hasPart property of MLO-AD. Possibilities to represent this relation should be analyzed and realized in future work.

The implemented prototype of the aggregator portal based on MLO-AD complies with following requirements of Section 3.3:

- The database for saving the content of learning opportunities and their providers is based on MLO-AD.
- The portal aggregates information about educational offers and institutions automatically and at regular intervals. It does not only update existing data, but also considers new information.
- A functionality for targeted searches is included.
- The prototype offers a social gadget as an integration into social networking systems.
- This gadget provides an interface with an input field to search for learning opportunities.
- Results of a search query are presented in a simple way.

The following two criteria have not been fully met; or rather they have been solved in a different way than was initially intended:

- The prototype does not harvest data about learning opportunities from two different providers. However, it offers learning opportunity providers a plug-in for the Sakai course management system and implements two different technologies to collect data from this system. This plug-in includes components for an interface for handling OAI-PMH requests, a generator for creating Atom feeds with MLO-AD, and a template with RDFa annotations for describing MLO-AD content. Those developments can easily be integrated with any existing Sakai 2.5.4.
- Information about educational offers is not directly collected from the training center of CRIM. The prototype uses examples of CRIM that are added to Sakai.

Besides the social gadget, the prototype offers a general web portal to search for learning opportunities. This portal additionally allows learning opportunity providers the possibility to add a link that refers to their OAI-PMH repository or Atom feed. Consequently, the aggregator system will automatically collect the data of this repository or feed at regular intervals.

Pending developments of the prototype particularly consider the security of the system. To ensure a good quality of the offered information about learning opportunities, the source and the provider of the information need to be identified. Therefore, it is essential that learning opportunity providers register and log in before they add a link that refers to their data. Additionally, only real educational offers and no spam data are allowed to be added to the database. Moreover, it still needs to manage the existing information about learning opportunities in the aggregator database. For example, educational offers must be removed from the database if they are not valid anymore.

Another necessary amelioration of the prototype is the precise adjustment of the data model to the structure of MLO-AD. Learning opportunities are already saved in the database based on the MLO-AD model. However, commonly used properties of the provider, specification, and instance are separately defined in each of them, and not in an abstract learning opportunity object as shown in Fig. 2.1 of Section 2.2. A refactoring of the data model and the code would make the latter clearer and more efficient. Moreover, each MLO-AD property in the prototype can have an occurrence of zero or one (except that the identifier has exactly one, and offers can have more than one). The specification of MLO-AD does not declare a maximum occurrence of its properties. A learning opportunity can have more values for the same property. This feature still must be implemented for the prototype.

The prototype is developed as an open source project under the Educational Community License, Version 1.0. Therefore, anyone who is interested in the project can contribute to the further development of this product. Many possible solutions and design suggestions described in this thesis have not been implemented yet. Furthermore, new ideas will come up due to the developments of MLO-AD, the Web, as well as other technologies. These ideas are, of course, more than welcome.

Appendix A

CD-ROM Content

Format: CD-ROM, Single Layer, ISO9660-Format

A.1 Diploma Thesis

This diploma thesis is included in the CD-ROM as PDF, DVI, and PostScript documents. The copyright of this thesis lies with its author.

Path: /thesis/

thesis.pdf	Diploma thesis as PDF file
thesis.dvi	Diploma thesis as DVI file
thesis.ps	Diploma thesis as PostScript file

All images of this thesis are provided as EPS files. Their copyright lies with the author of this thesis, unless otherwise stated.

Path: /thesis/images/

*.eps	Images of diploma thesis as EPS files

Online sources used in this thesis are saved on the CD-ROM as PDF documents. The copyright of these sources lies with the respective authors.

Path: /thesis/bibliography/

*.pdf	Copies of referenced online sources

A.2 Source Code

The CD-ROM also contains the implementation of the prototype. The source code and all necessary files for the installation are saved in the *Aggregator Portal* and *Sakai Plug-in* folders, whereas the latter provides divisions for the three different components of the plug-in.

Path: /prototype/aggregatorPortal/

MLOApp/	Aggregator portal created with GAE
INSTALL.txt	Information about the installation
mlo-gadget.xml	XML file of OpenSocial gadget

Path: /prototype/sakaiPlugin/AtomFeedGenerator/

atom-rdfa-src/	Source code of new Sakai component
INSTALL.txt	Information about the installation

Path: /prototype/sakaiPlugin/OAI-PMH/

oaicat-src/	Source code of modified OAICat
INSTALL.txt	Information about the installation
mlo_ad.xsd	Copy of Mark Stubbs' XML schema
mlo_example.xml	Example of an XML file containing MLO-AD
oaicat.jar	JAR File created from source code
oaicat.properties	Modified properties file for OAICat
oaicat.war	Copy of original OAICat WAR file

Path: /prototype/sakaiPlugin/RDFaTemplate/

reference/	Modified Sakai's reference component
site/	Modified Sakai's site component
velocity/	Modified Sakai's velocity component
INSTALL.txt	Information about the installation

Bibliography

[1] Adida, B. and M. Birbeck: *RDFa Primer: Bridging the Human and Data Web — W3C Working Group Note 14 October 2008.* World Wide Web Consortium, October 2008. http://www.w3.org/TR/xhtml-rdfa-primer.

[2] Alesso, H.P. and C.F. Smith: *Thinking on the Web: Berners-Lee, Gödel, and Turing.* John Wiley & Sons Inc., 2006.

[3] Allemang, D. and J. Hendler: *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL.* Morgan Kaufmann Publishers, 2008.

[4] Allsopp, J.: *Microformats: Empowering Your Markup for Web 2.0.* friends of ED, 2007.

[5] Beckett, D. and J. Broekstra: *SPARQL Query Results XML Format – W3C Recommendation 15 January 2008.* World Wide Web Consortium, January 2008. http://www.w3.org/TR/rdf-sparql-XMLres.

[6] Berners-Lee, T.: *Weaving the Web.* Harper Collins Publishers Inc., 1999.

[7] Berners-Lee, T., J. Hendler, and O. Lassila: *The Semantic Web.* Scientific American, 284:34–43, May 2001. http://www.sciam.com/article.cfm?id=the-semantic-web.

[8] Bouras, C., G. Kounenis, and I. Misedakis: *A Web Content Manipulation Technique Based on Page Fragmentation.* Journal of Network and Computer Applications, 30:563–585, 2007.

[9] Brickley, D. and R. Guha: *RDF Vocabulary Description Language 1.0: RDF Schema – W3C Recommendation 10 February 2004.* World Wide Web Consortium, February 2004. http://www.w3.org/TR/rdf-schema.

[10] Butow, E.: *User Interface Design for Mere Mortals.* For Mere Mortals Series. Addison-Wesley, 2007.

[11] Clark, K.G., L. Feigenbaum, and E. Torres: *SPARQL Protocol for RDF – W3C Recommendation 15 January 2008.* World Wide Web Consortium, January 2008. http://www.w3.org/TR/rdf-sparql-protocol.

[12] Cole, T.W. and M. Foulonneau: *Using the Open Archives Initiative Protocol for Metadata Harvesting.* Third Millennium Cataloging. Libraries Unlimited, 2007.

[13] Confederation of EU Rectors' Conferences and Association of European Universities (CRE): *The Bologna Declaration on the European Space for Higher Education: An Explanation.* The European Higher Education Area, February 2000. http://ec.europa.eu/education/policies/educ/bologna/bologna.pdf, includes the Joint Declaration of the European Ministers of Education convened in Bologna on the 19th of June 1999.

[14] David, D.M., L. Goertz, B. Hildebrandt, A. Janus, H.M.P. Sohn, U. Reichel, A. Reisky, L. Reß, C. Stracke, P. Tesch, and K. Unverricht: *PAS 1068 — Learning, Education, and Training with Special Consideration of e-Learning: Guideline for the Description of Educational Offers.* Beuth Verlag GmbH, December 2006.

[15] Graf, A.: *RDFa vs. Microformats: A Comparison on Inline Metadata Formats in (X)HTML.* Techn. rep., DERI — Digital Enterprise Research Institute Innsbruck, Technikerstraße 21a, A-6020 Innsbruck, Austria, April 2007.

[16] Haas, H. and A. Brown: *Web Services Glossary — W3C Working Group Note 11 February 2004.* World Wide Web Consortium, February 2004. http://www.w3.org/TR/ws-gloss.

[17] Hammersley, B.: *Developing Feeds with RSS and Atom.* O'Reilly, 2005.

[18] Harold, E.R. and W.S. Means: *XML in a Nutshell.* O'Reilly, 3rd ed., 2004.

[19] Heaton, J.: *Programming Spiders, Bots, and Aggregators in Java.* Sybex, 2002.

[20] Jenssen, A.: *ECTS Information Package/Course Catalogue with HTML- and XML-Bindings.* University of Oslo, October 2004. http://www.usit.uio.no/prosjekter/eSU/eSU-revisjon/CDM/cdm-ects.pdf.

[21] Johnson, D.: *RSS and Atom in Action: Web 2.0 Building Blocks.* Manning Publications Co., 2006.

[22] Kashyap, V., C. Bussler, and M. Moran: *The Semantic Web: Semantics for Data and Services on the Web.* Springer, 2008.

[23] King, R.: *The University in the Global Age.* Universities into the 21st Century. Palgrave Macmillan, 2004.

[24] Lagoze, C., H. Van de Sompel, M. Nelson, and S. Warner: *The Open Archives Initiative Protocol for Metadata Harvesting.* Open Archives Initiative, December 2008. http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm, Protocol Version 2.0 of 2002-06-14.

[25] Manola, F. and E. Miller: *RDF Primer – W3C Recommendation 10 February 2004.* World Wide Web Consortium, February 2004. http://www.w3.org/TR/rdf-primer.

[26] McGuinness, D.L. and F. van Harmelen: *OWL Web Ontology Language Overview – W3C Recommendation 10 February 2004.* World Wide Web Consortium, February 2004. http://www.w3.org/TR/owl-features.

[27] Metamatrix Development & Consulting AB: *EMIL — Education Information Markup Language*, 2004. http://www.elframework.org/projects/xcri/EMIL_PM%20v.1.0.pdf/download.

[28] Myllymaki, J.: *Effective Web Data Extraction with Standard XML Technologies.* Computer Networks, 39:635–644, 2002.

[29] Prud'hommeaux, E. and A. Seaborne: *SPARQL Query Language for RDF – W3C Recommendation 15 January 2008.* World Wide Web Consortium, January 2008. http://www.w3.org/TR/rdf-sparql-query.

[30] Pézeril, M.: *Course Description Metadata (CDM): A Relevant and Challenging Standard for Universities*, 2006. http://ariane.unil.ch/pdf/seminar/e-Quality_WS2_MPezeril_article.pdf.

[31] Ravaioli, S. and S. Velay: *Supporting the Bologna Process with IT Standards - New Emerging European Group for the Definition of Data Exchange Standards between Higher Education Institutions.* Kion, Unisolution, November 2007. http://www.moveonnet.eu/institutions/standards/downloads/rome20071109_finalreport.pdf, Report from the European workshop, 9th November 2007 in Rome.

[32] Reese, T. and K. Banerjee: *Building Digital Libraries.* No. 153 in *How-To-Do-It Manuals.* Neal-Schuman Publishers Inc., 2008.

[33] Richardson, L. and S. Ruby: *RESTful Web Services.* O'Reilly, 2007.

[34] Stubbs, M.: *XCRI (eXchanging Course-Related Information).* Techn. rep., Manchester Metropolitan University, Aytoun Street, Manchester, M1 3GH, November 2006.

[35] Studer, R., S. Grimm, and A. Abecker: *Semantic Web Services: Concepts, Technologies and Applications.* Springer, 2007.

[36] The Nielsen Company: *Global Faces and Networked Places.* Nielsen, March 2009. http://blog.nielsen.com/nielsenwire/wp-content/uploads/2009/03/nielsen_globalfaces_mar09.pdf, A Nielsen Report on Social Networking's New Global Footprint.

[37] Valentine, A. and R. Skeel: *The Business Case for the Development of a PESC Standard for an XML Format for Course Catalogs.* Letter of Intent of Postsecondary Electronic Standard Council, April 2005. http://www.pesc.org/library/docs/workgroups/course%20catalog%20workgroup/LetterOfIntent.pdf.

[38] Øverby, E., T. Hoel, and S. Wilson: *The History of MLO-AD.* Hypatia AS, June 2008. http://www.slideshare.net/toreh/mlo-history-2008-06-24.

[39] Øverby, E., C.M. Stracke, A. Heath, P. Karlberg, J. Pawlowski, T. Hoel, M. Collett, *et al.*: *Metadata for Learning Opportunities (MLO) — Advertising.* ISO/IEC JTC1 SC36 – Information Technology for Learning, Education, and Training, October 2008. http://www.percolab.com/gtn/images/b/b5/MLO-AD.pdf.

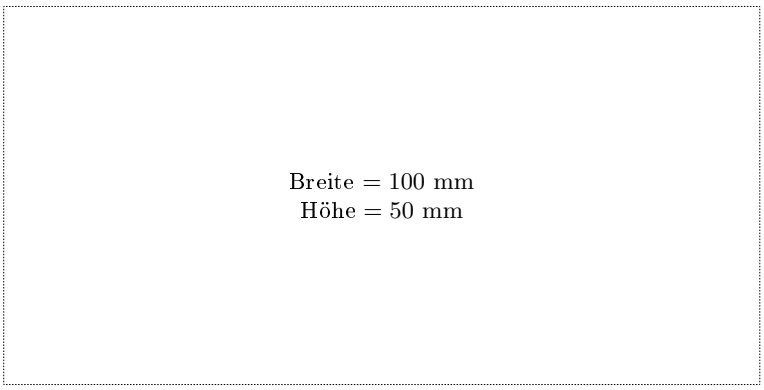
[40] Weerawarana, S., F. Curbera, F. Leymann, T. Storey, and D.F. Ferguson: *Web Services Platform Architecture: SOAP, WSDL, WS-Policy, WS-Addressing, WS-BPEL, WS-Reliable Messaging, and More.* Prentice Hall PTR, 2005.

[41] Witten, I.H., M. Gori, and T. Numerico: *Web Dragons: Inside the Myths of Search Engine Technology.* Morgan Kaufmann Publishers, 2007.

[42] Yee, R.: *Pro Web 2.0 Mashups: Remixing Data and Web Services.* The Expert's Voice in Web Development. Apress, 2008.

Messbox zur Druckkontrolle

— Druckgröße kontrollieren! —



— Diese Seite nach dem Druck entfernen! —

Publications du GTN-Québec

2012-03	*Soutien au développement de ressources numériques pour l'enseignement et l'apprentissage dans les universités québécoises – Rapport complet.* Rédigé par Line Cormier, Maureen Clapperton, Nicolas Gagnon, Michel Gendron, Robert Gérin-Lajoie et Jean Marcoux, 71 p.
2012-02	*Soutien au développement de ressources numériques pour l'enseignement et l'apprentissage dans les universités québécoises – Les faits saillants.* Rédigé par Line Cormier, Maureen Clapperton, Nicolas Gagnon, Michel Gendron, Robert Gérin-Lajoie et Jean Marcoux, 10 p.
2012-01	*Manuels de cours numériques – droit d'auteur et gestion, inventaire des solutions disponibles version 1.1.* Rédigé par Réjean Payette, 38 p.
2011-06	*Les tableaux numériques interactifs : considérations d'interopérabilité.* Rédigé par Marc-Antoine Parent, 28 p.
2011-05	*Fédération d'identité pour les organismes de l'éducation.* Rédigé par André Breton, 50 p.
2011-04	*Compte-rendu de participation, 26ème colloque annuel CSUN 2011.* Rédigé par Denis Boudreau, 14 p.
2011-03	*Les environnements d'apprentissage sont-ils en mutation ou en gestation?* Rédigé par Pierre-Julien Guay, Marcel Borduas, Yves Otis, Robet Paré et Sacha Leprêtre, 21 p.
2011-02	*Profil d'application québécois de métadonnées pour les opportunités d'étude, d'apprentissage et de formation (v.0.7.5)* Rédigé par Gilles Gauthier, 93 p.
2011-01	*Profil d'application Normetic 2.0 (v0.7.5)* Rédigé par Gilles Gauthier, 41 p.
2010-01	*Évaluation de fonctionnalités de traitement des métadonnées par Alfesco en comparaison avec Normetic.* Rédigé par François Vincent, 9 p.
2009-06	*Portrait des pratiques de sélection, de catalogage et de partage des documents numériques dans les bibliothèques.* Rédigé par Marie-Chantal Dufour, 48 p.
2009-05	*Accès aux contenus de formation en ligne : difficultés des apprenants handicapés et solutions pour assurer l'accessibilité des contenus.* Rédigé par Denis Boudreau, 21 p.
2009-04	*Développement MLO: Metadata for learning opportunities.* Rédigé par Olivier Gerbé et Thi-Lan-Anh Dinh, 32 p.
2009-03	*Concept and Prototype of an Aggregator Portal for Learning Opportunities Based on the MLO-AD Standard.* Rédigé par Katharina Bauer-Öppinger, 89 p.

(autres publications à la quatrième de couverture)

Publications du GTN-Québec (suite)

2009-02	*Identification des caractéristiques des modèles de diffusion de contenus numériques : recension des dépôts numériques existants – Partie 2.* Rédigé par Gabriel Dumouchel et Thierry Karsenti, 99 p.
2009-01	*Identification des caractéristiques des modèles de diffusion de contenus numériques : revue de littérature – Partie 1.* Rédigé par Gabriel Dumouchel et Thierry Karsenti, 54 p.
2008-05	*Ressources d'apprentissage et normes : la situation au Québec.* Rédigé par Christian Lafrance, 102 p.
2008-04	*Guide d'élaboration de fiches descriptives de ressources d'enseignement et d'apprentissage selon Normetic v1.2, profil d'application québécois du standard Learning Object Metadata (LOM).* Rédigé par Gérald Roberge, 57 p.
2008-03	*Profil d'application Normetic 1.2*. Rédigé par Gérald Roberge, 170 p.
2008-02	*Tableau du code XML à produire pour le vocabulaire de l'élément 5.2 de Normetic 1.2*. Rédigé par Gérald Roberge
2008-01	*Tableau du code XML à produire pour le vocabulaire de l'élément 5.6 de Normetic 1.2.* . Rédigé par Gérald Roberge
2007-01	*Portrait général des stratégies d'assurance qualité des ressources d'enseignement et d'apprentissage (REA) : à l'attention des gestionnaires.* Rédigé par Karin Lundgre-Cayrol, Suzanne Lapointe et Ileana De la Teja, 25 p.
2006-03	*Les normes, comment?* Rédigé par Gérald Roberge, 4 p.
2006-02	*Les normes, pourquoi?* Rédigé par Gérald Roberge, 4p.
2006-01	*Guide pour la sélection de REA.* Rédigé par Gérald Roberge, 10 p.
2005-01	*Le profil d'application Normetic, version 1.1.* Rédigé par Robert Thivierge, 8 p.
2003-01	*La description normalisée des ressources : vers un patrimoine éducatif – Normetic, version 1.0.* Sous la supervision de la CREPUQ et Novasys inc., 139 p.
2002-01	*Les normes et standards de la formation en ligne – État des lieux et enjeux.* Rédigé par Rachel Chouinard. Sous la supervision de la CREPUQ et du sous-comité SCTIC, 39 p.

Pour télécharger ces publications ou pour la liste complète des publications du GTN-Québec, voir le site Web **www.gtn-quebec.org/publications**